Adaptive Observers and Parametric Identification for Systems in Non-canonical Adaptive Observer Form

Ivan Y. Tyukin [a,d] Erik Steur [b] Henk Nijmeijer [c] Cees van Leeuwen [d]

[a] *Dept. of Mathematics, University of Leicester, Leicester, LE1 7RH, UK (Tel: +44-116-252-5106; I.Tyukin@le.ac.uk)*

[b] *Dept. of Mechanical Engineering, Eindhoven University of Technology, P.O. Box 513 5600 MB, Eindhoven, The Netherlands, (e.steur@tue.nl)*

[c] *Dept. of Mechanical Engineering, Eindhoven University of Technology, P.O. Box 513 5600 MB, Eindhoven, The Netherlands, (e-mail: h.nijmeijer@tue.nl)*

[d] *Laboratory for Perceptual Dynamics, RIKEN Brain Science Institute, Wako-shi, 351-0198, Japan, (ceesvl@brain.riken.jp)*

Abstract

We consider the problem of asymptotic reconstruction of the state and parameter values for dynamical systems that cannot be transformed into the canonical adaptive observer form. A solution to this problem is proposed for a class of systems for which the unknowns are allowed to be nonlinearly parameterized functions of state and time. Going beyond asymptotic Lyapunov stability, we provide for this class of systems a reconstruction technique, based on the concepts of weakly attracting sets, non-uniform convergence, and Poisson stability.

Key words: Adaptive observers, nonlinear parametrization, weakly attracting sets, unstable attractors, nonlinear systems, Poisson stability

1 Introduction

Often when studying and modeling real-life systems we come across a dynamical system of which we can register the input-output relations as functions of

* Corresponding author

time, but are unable to fully control the input the system receives over time. Suppose we know enough about the system to specify its equations, but to complete its description so that prediction of the system behavior over some interval of time is possible, we must recover the values of the internal system variables and parameters that are not available for direct observation. In this case, a powerful method for doing so is Adaptive Observer Design.

Observer-based methods can reconstruct both state variables and parameters, provided that the original system has, or can be transformed into, the *canonical adaptive observer form* [3]:

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{R}\mathbf{x} + \boldsymbol{\varphi}(y(t),t)\boldsymbol{\theta} + \mathbf{g}(t)\\ \mathbf{R} &= \begin{pmatrix} 0 & \mathbf{k}^T \\ 0 & \mathbf{F} \end{pmatrix}, \ \mathbf{x} = (x_0,\dots,x_n)^T \\ y(t) &= x_0(t)\end{aligned} \tag{1}$$

Function $\mathbf{g}:\mathbb{R}_{\geq 0}\rightarrow\mathbb{R}^{n+1}$ represents the *known* input, function $\boldsymbol{\varphi}:\mathbb{R}\times\mathbb{R}_{\geq 0}\rightarrow\mathbb{R}^{n+1}\times\mathbb{R}^d$ is a *known* function of the observed output $y(t)$, $\mathbf{k}=(k_1,\dots,k_n)$ is a vector of *known* constants, $\mathbf{F}$ is a *known* $n\times n$ matrix (usually diagonal), and $\boldsymbol{\theta}\in\mathbb{R}^d$ is a vector of *unknown* parameters. Algorithms are available in [3], [17] [1] that guarantee asymptotically stable (and sometimes even exponentially fast with arbitrary rate of convergence [19]) reconstruction of the unknowns, i.e. the state and parameter vectors.

However, many important systems in, for instance, chemical kinetics [7], [2] and biology [12], are not in canonical adaptive observer form. Unknown parameters in $\mathbf{F}$ and nonlinear parametrization characterize these systems. Absence of a canonical representation for these systems prevents straightforward application of standard observer-based techniques for state and parameter estimation.

A usual way to deal with this obstacle is to search for a bijective coordinate and parameter transformation reshaping the original system equations into the canonical form. Finding such a transformation, however, is a non-trivial problem. Available necessary and sufficient results in this direction apply to a too narrow class of systems (see e.g. [16]). Alternative, and possibly parameter-dependent, transformations may not be satisfactory either.

This leaves us with the undesirable options to use heuristic methods or exhaustive search. Whereas the former are not guaranteed to work, the latter can be forbiddingly expensive in terms of computational costs.

[1] See also [18] for the treatment of this problem when the regressor $\boldsymbol{\varphi}(y(t),t)\boldsymbol{\theta}$ in (1) is replaced with a nonlinearly parameterized function of the output y and parameter vector $\boldsymbol{\theta}$.

We propose a solution that combines the advantage of exponentially fast convergence with the flexibility of explorative behavior. Fast convergence is reserved for available estimators for state and linear parameters $\boldsymbol{\theta}$ of (1); exploration is reserved for the non-linear parameters an and the unknowns in $\mathbf{F}$. The proposed method therefore bears an overall similarity to [25], [11], [20], [21], [27]. However, instead of sticking to the framework of Lyaunov stability we relax the problem further and invoke the concepts of weakly attracting sets and relaxation times [22], [8], [9] in our argument. The method is based on results of [28] that allow us to analyze asymptotic convergence in nonlinear systems beyond the framework of Lyapunov design. We show that, subject to a condition of persistent excitation of the parameterized uncertainty, it is possible to reconstruct both linear and nonlinear parameters of the system. The latter result is illustrated with three examples.

The paper is organized as follows. In Section 2 we define the notation used throughout the paper. Section 3 provides the formal statement of the problem, Section 4 contains the main results, Section 5 provides illustrative examples, and Section 6 concludes with a brief discussion. The proofs of all formal statements are provided in the Appendix.

2 Notation

The following notational conventions are used throughout the paper:

- $\mathbb{R}_{>0} = \{x \in \mathbb{R} \mid x > 0\}$.
- $\mathbb{Z}$ denotes the set of integers, and $\mathbb{N}$ stands for the set of positive integers.
- The Euclidian norm of $\mathbf{x} \in \mathbb{R}^n$ is denoted by $\|\mathbf{x}\|$.
- By $L^n_\infty[t_0, T]$, $t_0 \geq 0$, $T \geq t_0$ we denote the space of all functions $\mathbf{f} : \mathbb{R}_{\geq 0} \to \mathbb{R}^n$ such that $\|\mathbf{f}\|_{\infty,[t_0,T]} = \operatorname{ess\,sup}\{\|\mathbf{f}(t)\|, t \in [t_0, T]\} < \infty$; $\|\mathbf{f}\|_{\infty,[t_0,T]}$ stands for the $L^n_\infty[t_0, T]$ norm of $\mathbf{f}(t)$.
- $\mathcal{C}^r$ denotes the space of continuous functions that are at least r times differentiable.
- Let $\mathcal{A}$ be a subset of $\mathbb{R}^n$, then for all $x \in \mathbb{R}^n$, we define $\operatorname{dist}(\mathcal{A}, x) = \inf_{q \in \mathcal{A}} \|x - q\|$.
- $\mathcal{O}(\cdot)$ denotes a function such that $\lim_{s \to 0} \mathcal{O}(s)/s = d$, $d \in \mathbb{R}$, $d \neq 0$.
- Finally, let $\epsilon \in \mathbb{R}_{>0}$, then $\|\mathbf{x}\|_\epsilon$ stands for the following:

$$\|\mathbf{x}\|_\epsilon = \begin{cases} \|\mathbf{x}\| - \epsilon, & \|\mathbf{x}\| > \epsilon, \\ 0, & \|\mathbf{x}\| \leq \epsilon. \end{cases}$$

3 Problem Formulation

We consider the following class of single-input-single-output nonlinear systems:

$$\begin{cases} \dot{x}_0 &= \theta_0^T\ \boldsymbol{\phi}_0(x_0,p_0,t) + \sum_{i=1}^n c_i(x_0,q_i,t)x_i + c_0(x_0,q_0,t) + \xi_0(t) + u(t) \\ \dot{x}_1 &= -\beta_1(x_0,\tau_1,t)\ x_1 + \theta_1^T\boldsymbol{\phi}_1(x_0,p_1,t) + \xi_1(t), \\ \quad\vdots & \\ \dot{x}_i &= -\beta_i(x_0,\tau_i,t)\ x_i + \theta_i^T\boldsymbol{\phi}_i(x_0,p_i,t) + \xi_i(t), \\ \quad\vdots & \\ \dot{x}_n &= -\beta_n(x_0,\tau_n,t)\ x_n + \theta_n^T\boldsymbol{\phi}_i(x_0,p_n,t) + \xi_n(t), \end{cases} \tag{2}$$

$$\begin{aligned} &y = (1,0,\ldots,0)^T\mathbf{x} = x_0,\ x_i(t_0) = x_{i,0}\in\mathbb{R}, \\ &\mathbf{x} = \mathrm{col}(x_0,x_1,\ldots,x_n),\ \theta_i = \mathrm{col}(\theta_{i,1},\ldots,\theta_{i,d_i}), \end{aligned}$$

where

$$\begin{aligned} \boldsymbol{\phi}_i &: \mathbb{R}\times\mathbb{R}^{m_i}\times\mathbb{R}_{\geq 0}\to\mathbb{R}^{d_i},\ \boldsymbol{\phi}_i\in\mathcal{C}^0,\ d_i,m_i\in\mathbb{N},\ i=\{0,\ldots,n\} \\ \beta_i &: \mathbb{R}\times\mathbb{R}\times\mathbb{R}_{\geq 0}\to\mathbb{R}_{>0},\ \beta_i\in\mathcal{C}^0,\ i=\{1,\ldots,n\} \\ c_i &: \mathbb{R}\times\mathbb{R}^{r_i}\times\mathbb{R}_{\geq 0}\to\mathbb{R},\ c_i\in\mathcal{C}^0,\ r_i\in\mathbb{N},\ i=\{0,\ldots,n\} \end{aligned}$$

are continuous and known functions, $u:\mathbb{R}_{\geq 0}\to\mathbb{R}$, $u\in\mathcal{C}^0$ is a known function of time modelling the control input, and $\xi_i:\mathbb{R}_{\geq 0}\to\mathbb{R}$, $\xi_i\in\mathcal{C}^0$ are functions that are unknown, yet bounded. The functions $\xi_i(t)$ represent *unmodeled dynamics* (e.g. noise).

Variable y in system (2) is the *output*, and the variables x_i, $i\geq 1$ are the components of state $\mathbf{x}$, that are not available for direct observation. Vectors $\theta_i\in\mathbb{R}^{d_i}$ consist of *linear* parameters of uncertainties in the right-hand side of the i-th equation in (2). Parameters $\tau_i\in\mathbb{R}$, $i=\{1,\ldots,n\}$ are the unknown parameters of time-varying relaxation rates, $\beta_i(x_0,\tau_i,t)$, of the state variables x_i, and vectors $p_i\in\mathbb{R}^{m_i}$, $q_i\in\mathbb{R}^{r_i}$, consist of the *nonlinear* parameters of the uncertainties. The functions $c_i(x_0,q_i,t)$ are supposed to be bounded.

Systems (2) model the dynamics of neural membranes [12]; they describe various chemical oscillators [2]; a number of important models in ecology, e.g. Lotka-Volterra equations, can be reduced to (2) by a, possibly parameter-dependent, change of coordinates (see our Examples section for details). For these reasons alone equations (2) constitute an interesting class of systems with remarkable practical relevance. On the other hand, and from a more theoretical viewpoint, system (2) may be viewed as a natural nonlinear extension

of the standard adaptive observer canonic form (1) in which the regressors are allowed to be nonlinearly parameterized and relaxation parameters, $\beta_i(\cdot)$, are time-varying and uncertain.

For notational convenience we denote:

$$\begin{aligned}
\boldsymbol{\theta} &= \mathrm{col}(\theta_0, \theta_1, \cdots, \theta_n), \\
\boldsymbol{\lambda} &= \mathrm{col}(p_0, q_0, \tau_1, p_1, q_1 \ldots, \tau_n, p_n, q_n), \\
s &= \dim(\boldsymbol{\lambda}) = n + \sum_{i=0}^{n}(m_i + r_i).
\end{aligned}$$

Symbols Ω_θ and Ω_λ, respectively, denote domains of admissible values for $\boldsymbol{\theta}$ and $\boldsymbol{\lambda}$.

The system state $\mathbf{x} = \mathrm{col}(x_0, x_1, \cdots, x_n)$ is not measured; only the values of the input $u(t)$ and the output $y(t) = x_0(t)$, $t \geq t_0$ in (2) are accessible over any time interval $[t_0, t]$ that belongs to the history of the system. The actual values of parameters $\boldsymbol{\theta}$, $\boldsymbol{\lambda}$ are assumed to be unknown *a-priori*. For simplicity of the proofs, we assume them to lay within a hypercube with known bounds: $\theta_{i,j} \in [\theta_{i,\min}, \theta_{i,\max}]$, $\lambda_i \in [\lambda_{i,\min}, \lambda_{i,\max}]$.

To reconstruct the unknown state and parameters of system from the values of $y(t) = x_0(t)$, $u(t)$, we should find an auxiliary system

$$\dot{\mathbf{q}} = \mathbf{f}(\mathbf{q}, y(t), t), \ \ \mathbf{q}(t_0) = \mathbf{q}_0, \ \ \mathbf{q} \in \mathbb{R}^q \tag{3}$$

such that for some given $\mathbf{h} : \mathbb{R}^q \to \mathbb{R}^p$, $p = \dim(\mathrm{col}(\boldsymbol{\theta}^T, \boldsymbol{\lambda}^T, \mathbf{x}^T))$, $\mathbf{q}_0$, $\delta \in \mathbb{R}_{>0}$ and all $t_0 \in \mathbb{R}_{\geq 0}$ the following property holds:

$$\exists\ t' \geq t_0 : \|\mathbf{h}(\mathbf{q}(t, \mathbf{q}_0)) - \mathbf{w}(t)\| < \delta, \ \forall\ t \geq t' \tag{4}$$

where $\mathbf{w}(t) = \mathrm{col}(\boldsymbol{\theta}, \boldsymbol{\lambda}, \mathbf{x}(t, \mathbf{x}_0))$. System (3) is usually referred to as an *adaptive observer.*

The goal requirement for our adaptive observer is stated in the form of inequality (4), instead of the more usual requirement $\lim_{t\to\infty} \mathbf{h}(\mathbf{q}(t, \mathbf{q}_0)) - \mathbf{w}(t) = 0$. This is because we allow unmodeled dynamics, $\xi_i(t)$, in the right-hand side of (2). As a result of such practically important addition, there is a possibility that (4) is satisfied not only for the original system exclusively, but it may also hold for a set of systems. The vector-fields in the right hand side of each element of this set are to be sufficiently close to that of (2), yet their parameters could be different. Instead of just one value of unknown parameter vectors $\boldsymbol{\theta}$, $\boldsymbol{\lambda}$ we therefore have to deal with a set of $\boldsymbol{\theta}$, $\boldsymbol{\lambda}$ corresponding to the solutions of (2) that over time are sufficiently close. This set of model parameters is referred to as an *equivalence class* of (2).

In the next section we demonstrate that despite the non-uniqueness of the solution, asymptotic reconstruction of state and parameters of (2) is achievable, subject to a standard persistency of excitation condition.

4 Main Results

Traditionally, methods of observer design are based on the notion of Lyapunov stability of solutions. If a globally asymptotically Lyapunov-stable solution exists in the state space of the combined system (2), (3) such that the requirements (4) are satisfied, then (3) is the observer needed. However, as we mentioned earlier, in systems with nonlinear parameterization and unmodeled perturbations distinct parameter values in the right-hand side of (2) may correspond to indistinguishable input-output behavior. Hence multiple distinct invariant sets in the extended state space will co-exist, and only local stability would generally be guaranteed. In order to deal with these issues, novel concepts of observer design are needed.

We propose a method for finding adaptive observers (3) for (2) that rests on a two-fold idea. First, instead of the standard requirement that the dynamics of errors of state and parameter estimates is globally asymptotically stable in the sense of Lyapunov, we will require mere convergence of the estimates as specified by (4). This relaxation allows us to invoke a wider range of tools for accessing convergence such as in [28].

Second, similarly to canonical observer schemes [16], [17], [18], [19] we shall be relying on the possibility to evaluate the integrals

$$\mu_i(t, \tau_i, p_i) \triangleq \int_{t_0}^{t} e^{-\int_{\tau}^{t} \beta_i(x_0(\chi), \tau_i, \chi) d\chi} \boldsymbol{\phi}_i(x_0(\tau), p_i, \tau) d\tau \quad (5)$$

at a given time t and for the given values of τ_i, p_i within a given accuracy. In classical adaptive observer schemes, the values of $\beta_i(x_0, \tau_i, t)$ are constant. This allows us to transform the original equations by a (possibly parameter-dependent) non-singular linear coordinate transformation, $\Phi : \mathbf{x} \mapsto \mathbf{z}$, $x_1 = z_1$, into an equivalent form in which the values of all time constants are known. In the new coordinates the variables $z_2, \ldots, z_n$ can be estimated by integrals (5) in which the values of $\beta_i(x_0, \tau_i, t)$ are constant and known. This is usually done by using auxiliary linear filters. In our case, the values of $\beta_i(x_0, \tau_i, t)$ are not constant and are unknown due to the presence of τ_i. Yet if the values of τ_i would be known we could still estimate the values of integrals (5) as follows

$$\int_{t_0}^{t} e^{-\int_{\tau}^{t} \beta_i(x_0(\chi),\tau_i,\chi)d\chi} \boldsymbol{\phi}_i(x_0(\tau),p_i,\tau)d\tau \simeq \\ \int_{t-T}^{t} e^{-\int_{\tau}^{t} \beta_i(x_0(\chi),\tau_i,\chi)d\chi} \boldsymbol{\phi}_i(x_0(\tau),p_i,\tau)d\tau \triangleq \bar{\mu}_i(t,\tau_i,p_i), \tag{6}$$

where $T \in \mathbb{R}_{>0}$ is sufficiently large and $t \geq T + t_0$.

Alternatively, if $\boldsymbol{\phi}_i(x_0(t),p_i,t)$, $\beta_i(x_0(t),\tau_i,t)$ are periodic with rationally - dependent periods and satisfy the Dini condition in t, integrals (5) can be estimated invoking a Fourier expansion. Notice that for continuous and Lipschitz in p_i functions $\mu_i(t,\tau_i,p_i)$ the coefficients of their Fourier expansion remain continuous and Lipschitz with respect to p_i.

In what follows we show that the availability of a computational scheme which is capable of evaluating the integrals (5) with sufficient accuracy enables us to design an adaptive observer that, under some mild and standard persistency of excitation assumptions, ensures asymptotic estimation of the unknown parameter values of (2).

4.1 Observer definition and assumptions

Consider the following function $\boldsymbol{\varphi}(x_0,\boldsymbol{\lambda},t) : \mathbb{R} \times \mathbb{R}^s \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}^d$, $d = \sum_{i=0}^{n} d_i$:

$$\begin{aligned}\boldsymbol{\varphi}(x_0,\boldsymbol{\lambda},t) = (\boldsymbol{\phi}_0(x_0,p_0,t), c_1(x_0,q_1,t)\mu_1(t,\tau_1,p_1),\ldots \\ \ldots, c_n(x_0,q_n,t)\mu_n(t,\tau_n,p_n))^T\end{aligned} \tag{7}$$

The function $\boldsymbol{\varphi}(x_0,\boldsymbol{\lambda},t)$ is a concatenation of $\boldsymbol{\phi}_0(\cdot)$ and integrals (5). We assume that the values of $\boldsymbol{\varphi}(x_0,\boldsymbol{\lambda},t)$ can be efficiently estimated for all x_0, $\boldsymbol{\lambda}$, $t \geq 0$ up to a small mismatch. In other words, we suppose that there exists a function $\bar{\boldsymbol{\varphi}}(x_0,\boldsymbol{\lambda},t)$ such that the following property holds:

$$\|\bar{\boldsymbol{\varphi}}(x_0,\boldsymbol{\lambda},t) - \boldsymbol{\varphi}(x_0,\boldsymbol{\lambda},t)\| \leq \Delta_{\varphi},\ \ \Delta_{\varphi} \in \mathbb{R}_{>0}, \tag{8}$$

where values of $\bar{\boldsymbol{\varphi}}(x_0,\boldsymbol{\lambda},t)$ are efficiently computable for all x_0, $\boldsymbol{\lambda}$, t (see e.g. (6) for an example of such approximations), and Δ_{φ} is sufficiently small.

If parameters τ_i, p_i, and q_i in the right-hand side of (2) would be known and $c_i(x_0,q_i,t) = 1$, $\beta_i(x_0,\tau_i,t) = \tau_i$, then the function $\boldsymbol{\varphi}(x_0,\boldsymbol{\lambda},t)$ could be estimated by $(\boldsymbol{\phi}_0(x_0,t),\eta_1,\ldots,\eta_n)$ where η_i are the solutions of the following auxiliary system (filter)

$$\dot{\eta}_i = -\tau_i\eta_i + \boldsymbol{\phi}_i(x_0,p_i,t) \tag{9}$$

with zero initial conditions. Systems like (9) are inherent components of standard adaptive observers [13], [17]. In our case we suppose that the values of τ_i,

q_i, p_i are not know a-priori and that $c_i(x_0, q_i, t)$, $\beta_i(x_0, \tau_i, t)$ are not constant. Therefore, we replace η_i with their approximations, e.g. as in (6):

$$\begin{aligned}\bar{\boldsymbol{\varphi}}(x_0, \boldsymbol{\lambda}, t) = (&\boldsymbol{\phi}_0(x_0, p_0, t), c_1(x_0, q_1, t)\bar{\mu}_1(t, \tau_1, p_1), \ldots \\ &\ldots, c_n(x_0, q_n, t)\bar{\mu}_n(t, \tau_n, p_n))^T .\end{aligned}$$

For periodic $\boldsymbol{\phi}_i(x_0(t), p_i, t)$, $\beta_i(x_0(t), \tau_i, t)$ a Fourier expansion can be employed to define $\bar{\boldsymbol{\varphi}}(x_0, \boldsymbol{\lambda}, t)$. The value of Δ_φ in (8) stands for the accuracy of approximation, and as a rule of thumb the more computational resources are devoted to approximate $\boldsymbol{\varphi}(x_0, \boldsymbol{\lambda}, t)$ the smaller is the value of Δ_φ.

With regard to the functions $\xi_i(t)$ in (2) we suppose that an upper bound, Δ_ξ, of the following sum is available:

$$\sum_{i=1}^{n} \frac{1}{\tau_i} \|\xi_i(\tau)\|_{\infty,[t_0,\infty]} + \|\xi_0(\tau)\|_{\infty,[t_0,\infty]} \leq \Delta_\xi,\ \Delta_\xi \in \mathbb{R}_{\geq 0}. \tag{10}$$

Denoting $c_0(x_0, q_0, t) = c_0(x_0, \boldsymbol{\lambda}, t)$, for notational convenience, we can now define the observer as

$$\left\{\begin{array}{l} \dot{\hat{x}}_0 = -\alpha(\hat{x}_0 - x_0) + \hat{\boldsymbol{\theta}}^T \bar{\boldsymbol{\varphi}}(x_0, \hat{\boldsymbol{\lambda}}, t) + c_0(x_0, \hat{\boldsymbol{\lambda}}, t) + u(t) \\ \dot{\hat{\boldsymbol{\theta}}} = -\gamma_\theta(\hat{x}_0 - x_0)\bar{\boldsymbol{\varphi}}(x_0, \hat{\boldsymbol{\lambda}}, t),\ \gamma_\theta, \alpha \in \mathbb{R}_{>0} \end{array}\right. \tag{11}$$

$$\dot{\hat{x}}_i = -\beta_i(x_0, \hat{\tau}_i, t)\hat{x}_i + \hat{\theta}_i^T \boldsymbol{\phi}_i(x_0, \hat{p}_i, t),\ i = \{1, \ldots, n\}, \tag{12}$$

where $\hat{\boldsymbol{\theta}} = \mathrm{col}(\hat{\theta}_0, \hat{\theta}_1, \cdots, \hat{\theta}_n)$ is the vector of estimates of $\boldsymbol{\theta}$. The components of vector $\hat{\boldsymbol{\lambda}} = \mathrm{col}(\hat{p}_0, \hat{q}_0, \hat{\tau}_1, \hat{p}_1, \hat{q}_1, \ldots, \hat{\tau}_n, \hat{p}_n, \hat{q}_n) = \mathrm{col}(\hat{\lambda}_1, \ldots, \hat{\lambda}_s)$, with $s = \dim(\boldsymbol{\lambda})$, evolve according to the following equations

$$\left\{\begin{array}{ll} \dot{\hat{x}}_{1,j} & = \gamma \cdot \omega_j \cdot e \cdot \left(\hat{x}_{1,j} - \hat{x}_{2,j} - \hat{x}_{1,j}\left(\hat{x}_{1,j}^2 + \hat{x}_{2,j}^2\right)\right) \\ \dot{\hat{x}}_{2,j} & = \gamma \cdot \omega_j \cdot e \cdot \left(\hat{x}_{1,j} + \hat{x}_{2,j} - \hat{x}_{2,j}\left(\hat{x}_{1,j}^2 + \hat{x}_{2,j}^2\right)\right) \\ \hat{\lambda}_j(\hat{x}_{1,j}) & = \lambda_{j,\min} + \frac{\lambda_{j,\max} - \lambda_{j,\min}}{2}(\hat{x}_{1,j} + 1), \\ e & = \sigma(\|x_0 - \hat{x}_0\|_\varepsilon), \end{array}\right. \tag{13}$$

$$j = \{1, \ldots, s\},\ \hat{x}_{1,j}^2(t_0) + \hat{x}_{2,j}^2(t_0) = 1, \tag{14}$$

where $\sigma(\cdot) : \mathbb{R} \to \mathbb{R}_{\geq 0}$ is a bounded continuous function, i.e. $\sigma(\upsilon) \leq S \in \mathbb{R}_{>0}$, and $|\sigma(\upsilon)| \leq |\upsilon|$ for all $\upsilon \in \mathbb{R}$. We set $\omega_j \in \mathbb{R}_{>0}$ and let ω_j be *rationally-independent*:

$$\sum \omega_j k_j \neq 0,\ \forall\ k_j \in \mathbb{Z}. \tag{15}$$

A proof that system (11) – (13) can serve as the desired observer (3) requires the concepts of non-uniform convergence [22], [9], non-uniform small-gain theorems [28], and the notions of λ-*uniform persistency of excitation* [14] and *nonlinear persistency of excitation* [6]:

Definition 1 (λ-uniform persistency of excitation) *Let $\boldsymbol{\varphi}: \mathbb{R}_{\geq 0} \times \mathcal{D} \to \mathbb{R}^{n \times m}$, $\mathcal{D} \subset \mathbb{R}^s$ be a continuous function. We say that $\boldsymbol{\varphi}(t, \boldsymbol{\lambda})$ is λ-uniformly persistently exciting (λ-uPE) if there exist $\mu \in \mathbb{R}_{>0}$, $L \in \mathbb{R}_{>0}$ such that for each $\boldsymbol{\lambda} \in \mathcal{D}$*

$$\int_t^{t+L} \boldsymbol{\varphi}(t, \boldsymbol{\lambda}) \boldsymbol{\varphi}(t, \boldsymbol{\lambda})^T d\tau \geq \mu I \ \forall t \geq t_0. \tag{16}$$

In contrast to conventional definitions [26], the present notion requires that the lower bound for the integral $\int_t^{t+L} \boldsymbol{\varphi}(t, \boldsymbol{\lambda}) \boldsymbol{\varphi}(t, \boldsymbol{\lambda})^T d\tau$ in (16) does not vanish for all $\boldsymbol{\lambda} \in \mathcal{D}$, and is separated away from zero. We need this property in order to determine the linear parts, θ_i, of the parametric uncertainties in model (2).

To reconstruct the nonlinear part of the uncertainties, $\boldsymbol{\lambda}$, we will require that $\bar{\boldsymbol{\varphi}}(x_0, \boldsymbol{\lambda}, t)$ is nonlinearly persistently exciting in $\boldsymbol{\lambda}$. Here we adopt the definition of nonlinear persistent excitation from [6] with a minor modification. The modification is needed to account for a possibility that

$$\bar{\boldsymbol{\varphi}}(x_0, \boldsymbol{\lambda}, t) = \bar{\boldsymbol{\varphi}}(x_0, \boldsymbol{\lambda}', t), \ \boldsymbol{\lambda} \neq \boldsymbol{\lambda}', \ t \in \mathbb{R},$$

which is the case, for example if $\bar{\boldsymbol{\varphi}}(x_0, \boldsymbol{\lambda}, t)$ is periodic in $\boldsymbol{\lambda}$. The modified notion is presented in Definition 2 below.

Definition 2 (Nonlinear persistency of excitation) *The function $\bar{\boldsymbol{\varphi}}(x_0, \boldsymbol{\lambda}, t)$ is nonlinearly persistently exciting if there exist $L, \beta \in \mathbb{R}_{>0}$ such that for all $\boldsymbol{\lambda}, \boldsymbol{\lambda}' \in \Omega_\lambda$ and $t \in \mathbb{R}$ there exists $t' \in [t - L, t]$ ensuring that the following inequality holds*

$$\|\bar{\boldsymbol{\varphi}}(x_0, \boldsymbol{\lambda}, t) - \bar{\boldsymbol{\varphi}}(x_0, \boldsymbol{\lambda}', t')\| \geq \beta \cdot \ \mathrm{dist}(\mathcal{E}(\lambda), \boldsymbol{\lambda}'), \tag{17}$$

$$\mathcal{E}(\boldsymbol{\lambda}) = \{\boldsymbol{\lambda}' \in \Omega_\lambda | \ \bar{\boldsymbol{\varphi}}(x_0, \boldsymbol{\lambda}', t) = \bar{\boldsymbol{\varphi}}(x_0, \boldsymbol{\lambda}, t) \ \forall \ t \in \mathbb{R}\} \tag{18}$$

The symbol $\mathcal{E}(\boldsymbol{\lambda})$ denotes the equivalence class for $\boldsymbol{\lambda}$, and $\mathrm{dist}(\mathcal{E}(\lambda), \boldsymbol{\lambda}')$ in (17) substitutes the Euclidian norm in the [6] original definition. The nonlinear persistency of excitation condition (17) is very similar to its linear counterpart (16). In fact (16) can be written in the form of inequality (17), cf. [23]. For further discussion of these notions, see [6], [14].

In our work we use condition (16) to ensure state boundedness of the observer and to enable reconstruction of linear parameters θ_i in the original equations (2). Condition (17) is used to formulate the requirements for successful reconstruction of nonlinear parameters p_i and unknown relaxation times τ_i.

4.2 Asymptotic properties of the observer

The main results of this section are provided in Theorems 3 and 5. Theorem 3 establishes conditions for state boundedness of the observer, and states its general asymptotic properties. Theorem 5 specifies a set of conditions for the possibility of asymptotic reconstruction of θ_i, τ_i, and p_i, up to their equivalence classes and small mismatch due to errors.

Proofs of Theorems 3, 5 and other auxiliary results are provided in Appendix.

Theorem 3 (Boundedness) *Let system (2), (11) – (13) be given. Assume that function $\bar{\boldsymbol{\varphi}}(x_0(t), \boldsymbol{\lambda}, t)$ is λ-uniformly persistently exciting, and the functions $\bar{\boldsymbol{\varphi}}(x_0(t), \boldsymbol{\lambda}, t)$, $c_0(x_0(t), \boldsymbol{\lambda}, t)$ are Lipschitz in $\boldsymbol{\lambda}$:*

$$\begin{aligned} \|\bar{\boldsymbol{\varphi}}(x_0(t), \boldsymbol{\lambda}, t) - \bar{\boldsymbol{\varphi}}(x_0(t), \boldsymbol{\lambda}', t)\| &\leq D\|\boldsymbol{\lambda} - \boldsymbol{\lambda}'\|, \\ \|c_0(x_0(t), \boldsymbol{\lambda}, t) - c_0(x_0(t), \boldsymbol{\lambda}', t)\| &\leq D_c\|\boldsymbol{\lambda} - \boldsymbol{\lambda}'\|. \end{aligned} \tag{19}$$

Then there exist numbers $\varepsilon > 0$, $\gamma^* > 0$ such that for all $\gamma \in (0, \gamma^*]$:

1) trajectories of the closed loop system (11) – (13) are bounded and

$$\lim_{t\to\infty} \|\hat{x}_0(t) - x_0(t)\|_\varepsilon = 0; \tag{20}$$

2) there exists $\boldsymbol{\lambda}^* \in \Omega_\lambda$, $\kappa \in \mathbb{R}_{>0}$ such that

$$\begin{aligned} \lim_{t\to\infty} \hat{\boldsymbol{\lambda}}(t) &= \boldsymbol{\lambda}^* \\ \limsup_{t\to\infty} \|\hat{\boldsymbol{\theta}}(t) - \boldsymbol{\theta}\| &< \kappa((D\|\boldsymbol{\theta}\| + D_c)\|\boldsymbol{\lambda}^* - \boldsymbol{\lambda}\| + 2\Delta). \end{aligned} \tag{21}$$

$$\Delta = \|\boldsymbol{\theta}\|\Delta_\varphi + \Delta_\xi \tag{22}$$

Remark 4 Theorem 3 assures that the estimates $\hat{\boldsymbol{\theta}}(t)$, $\hat{\boldsymbol{\lambda}}(t)$ asymptotically converge to a neighborhood of the actual values $\boldsymbol{\theta}$, $\boldsymbol{\lambda}$. It does not specify, however, how close these estimates are to the true values of $\boldsymbol{\theta}$, $\boldsymbol{\lambda}$.

Our next result states that if the values of Δ_φ and Δ_ξ:

$$\begin{aligned} &\|\boldsymbol{\varphi}(x_0, \boldsymbol{\lambda}, t) - \bar{\boldsymbol{\varphi}}(x_0, \boldsymbol{\lambda}, t)\| \leq \Delta_\varphi \\ &\sum_{i=1}^{n} \frac{1}{\tau_i}\|\xi_i(\tau)\|_{\infty,[t_0,\infty]} + \|\xi_0(\tau)\|_{\infty,[t_0,\infty]} \leq \Delta_\xi \end{aligned}$$

in (8), (10) are small, e.g. $\bar{\boldsymbol{\varphi}}(x_0(t), \boldsymbol{\lambda}, t)$ approximates $\boldsymbol{\varphi}(x_0(t), \boldsymbol{\lambda}, t)$ with sufficiently high accuracy and the unmodeled dynamic is negligible, the estimates $\hat{\boldsymbol{\theta}}(t)$, $\hat{\boldsymbol{\lambda}}(t)$ will converge to small neighborhoods of the equivalence classes of $\boldsymbol{\theta}$, $\boldsymbol{\lambda}$. The sizes of these neighborhoods are shown to be bounded from above

by monotone functions Δ:

$$\Delta = \|\boldsymbol{\theta}\|\Delta_{\varphi} + \Delta_{\xi}$$

vanishing at zero. Formally this result is stated in Theorem 5 below

Theorem 5 (Convergence) *Let the assumptions of Theorem 3 hold, assume that $\bar{\varphi}_0(x_0, \boldsymbol{\lambda}, t) \in \mathcal{C}^1$, the derivative $\partial\bar{\boldsymbol{\varphi}}_0(x_0(t), \boldsymbol{\lambda}, t)/\partial t$ is globally bounded, and $\Delta = \|\boldsymbol{\theta}\|\Delta_{\varphi} + \Delta_{\xi}$ is small[2]. Then there exist numbers $\varepsilon > 0$, $\gamma^* > 0$ such that for all $\gamma_i \in (0, \gamma^*)$*

1)

$$\limsup_{t\to\infty} \|\hat{\boldsymbol{\theta}}(t) - \boldsymbol{\theta}\| = \mathcal{O}(\sqrt{\Delta}) + \mathcal{O}(\Delta)$$

2) in case $\boldsymbol{\theta}^T\bar{\boldsymbol{\varphi}}(x_0(t), \boldsymbol{\lambda}, t) + c_0(x_0(t), \boldsymbol{\lambda}, t)$ is nonlinearly persistently exciting with respect to $\boldsymbol{\lambda}$, then the estimates $\hat{\boldsymbol{\lambda}}(t)$ converge into a small vicinity of $\mathcal{E}(\lambda)$:

$$\limsup_{t\to\infty}\ \mathrm{dist}(\hat{\boldsymbol{\lambda}}(t), \mathcal{E}(\boldsymbol{\lambda})) = \mathcal{O}(\sqrt{\Delta}) + \mathcal{O}(\Delta) \tag{23}$$

5 Examples

We illustrate our method with three examples. The first example deals with the problem of parameter reconstruction of the Duffing oscillator assuming that only one state variable is available for direct observation. This problem received considerable attention in the past and hence can serve as a benchmark. Existence of an output feedback steering the state of the Duffing system to an arbitrary small neighborhood of the origin in the presence of parametric uncertainties was demonstrated in [24], [15]. Recently, a method for asymptotic reconstruction of all parameters of the Duffing system was published in [4]. Here we show how our own method performs in this task provided that the frequency of external perturbations is known.

The second and the third examples address a more difficult class of problems in which the uncertain parameters enter the model nonlinearly and/or the functions $c_i(x_0, q_i, t)$ in (2) are not constant. In particular, we consider a two-component bio-reactor model [2] and a bilinear Lotka-Volterra system. We show that in both cases, subject to the chosen tolerance margin, state and parameters of these models can be successfully estimated from input-output observations with our adaptive observer.

[2] The bounds defining smallness of Δ are provided in (A.35)–(A.37) in Appendix

Example 1. Parameter estimation of the Duffing oscillator. Consider the following system of differential equations

$$\begin{aligned} \dot{x}_0 &= x_1 \\ \dot{x}_1 &= -\delta x_1 - \beta x_0 - \alpha x_0^3 + \gamma \cos(\omega t),\ y = x_0, \end{aligned} \tag{24}$$

where $\delta \in (0.1, 1)$, β, α, $\gamma \in \mathbb{R}$, are parameters of which the values are supposed to be unknown, and x_0 measured. The value the frequency of periodic forcing, i.e. ω, is supposed to be known. In this example we set

$$\delta = 0.2,\ \alpha = 1,\ \beta = -1,\ \gamma = 0.3,\ \omega = 1, \tag{25}$$

and the task is to reconstruct the values of δ, α, β, and γ asymptotically from the measurements of x_0.

For convenience we rewrite (24) using the notational agreements as in (2):

$$\begin{aligned} \dot{x}_0 &= x_1 \\ \dot{x}_1 &= -\tau_1 x_1 + \theta_{1,1} x_0 + \theta_{1,2} x_0^3 + \theta_{1,3} \cos(\omega t). \end{aligned} \tag{26}$$

According to (7), we start with constructing the following regressor function

$$\begin{aligned} \boldsymbol{\varphi}(x_0, \boldsymbol{\lambda}, t) &= (\varphi_1(x_0, \boldsymbol{\lambda}, t), \varphi_2(x_0, \boldsymbol{\lambda}, t), \varphi_3(x_0, \boldsymbol{\lambda}, t)) \\ &= \left(\int_{t_0}^t e^{-\tau_1(t-\tau)} x_0(\tau) d\tau, \int_{t_0}^t e^{-\tau_1(t-\tau)} x_0^3(\tau) d\tau, \int_{t_0}^t e^{-\tau_1(t-\tau)} \cos(\omega\tau) d\tau \right), \end{aligned}$$

where $\boldsymbol{\lambda} = \tau_1$. For the given parametrization of (24) the regressor satisfies the requirement of uniform persistency of excitation (the value of μ ranges from 1.2 for $\tau_1 = 1$ to 1000 for $\tau_1 = 0.2$ at $L = 500$). Following (11) – (13), the structure of the observer is defined as follows:

$$\begin{cases} \dot{\hat{x}}_0 = -(\hat{x}_0 - x_0) + \sum_{i=1}^{3} \hat{\theta}_{1,i} \varphi_i(x_0, \boldsymbol{\lambda}, t) \\ \dot{\hat{\theta}}_{1,i} = -\gamma_\theta (\hat{x}_0 - x_0) \varphi_i(x_0, \boldsymbol{\lambda}, t) \end{cases} \tag{27}$$

$$\begin{cases} \dot{\hat{x}}_{1,1} &= \gamma \cdot e \cdot \left(\hat{x}_{1,1} - \hat{x}_{2,1} - \hat{x}_{1,1} \left(\hat{x}_{1,1}^2 + \hat{x}_{2,1}^2 \right) \right) \\ \dot{\hat{x}}_{2,1} &= \gamma \cdot e \cdot \left(\hat{x}_{1,1} + \hat{x}_{2,1} - \hat{x}_{2,1} \left(\hat{x}_{1,1}^2 + \hat{x}_{2,1}^2 \right) \right) \\ \hat{\lambda}(\hat{x}_{1,1}) &= \lambda_{\min} + \frac{\lambda_{\max} - \lambda_{\min}}{2} (\hat{x}_{1,1} + 1), \\ e &= \sigma(\| x_0 - \hat{x}_0 \|_\varepsilon), \end{cases} \tag{28}$$

where $\sigma(\cdot) = \tanh(\cdot)$ and the parameters of the observer are set as follows: $\gamma = 0.2$, $\gamma_\theta = 2$, $\lambda_{\min} = 0.1$, $\lambda_{\max} = 1.1$, $\varepsilon = 0.01$. Performance of adaptive observer system (26), (27), (28) is illustrated with Fig. 1. In this example the estimates reached 5-percent neighborhood of the true values of model parameters at 2000 units of time. To achieve the same accuracy by the exhaustive-search

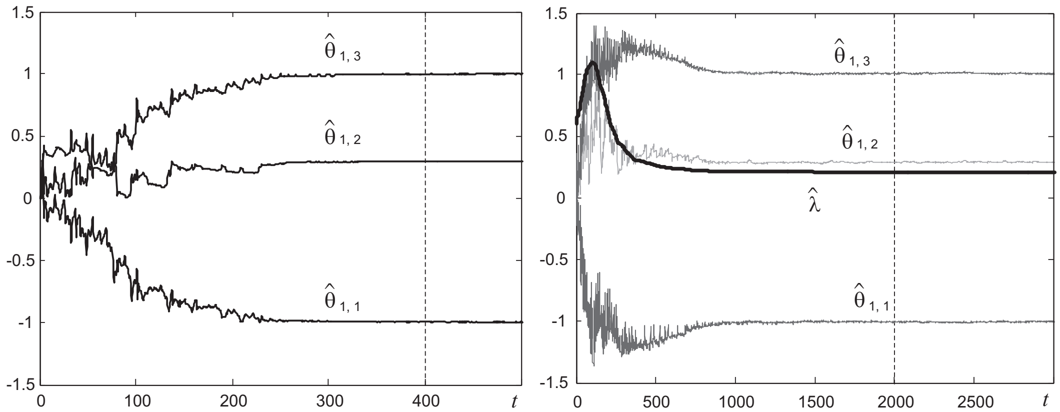

Fig. 1. Left: estimates $\hat{\theta}_i$ as functions of time for fixed $\hat{\lambda} = \tau_1$. We see that $\hat{\theta}_i$ converge to a small vicinity (5%) of their true value in 400 sec of model time. Right: estimates $\hat{\theta}_i$ and $\hat{\lambda}$ as functions of time in the closed-loop system (26), (27), (28). All estimates converge to small neighborhoods of their true value in 2000 units of time.

would require covering the interval $[0.1, 1.1]$ with a 0.01-resolution grid, for which 100 tests are needed. Thus the overall time spent to recover the value of τ_1 would amount to 40000 units of time, which is already in this simple case much larger than the time that our method requires.

Example 2. A Bio-reactor model. Consider the following system of equations [2]:

$$\begin{aligned} \dot{s}_0 &= -ds_0 + du(t) - br(s_0,k)s_1 \\ \dot{s}_1 &= -ds_1 + r(s_0,k)s_1,\ \ r(s_0,k) = \frac{r_{\max}s_0}{s_0+k}, \\ y &= s_0, \end{aligned} \tag{29}$$

where s_0, s_1 are products (concentrations), d is the dilution rate, $r(s_0,k)$ is a Monod-type nonlinearity, $r_{\max}$ is the maximum growth rate, $b > 0$ is the growth yield, and k is the “half saturation” constant. The measured output is the variable s_0. The values of parameters $r_{\max}$, b are supposed to be known, and precise values of parameters d, k are considered unknown.

According to the requirements of our method we need that the system equations are presented in the form (2) where $\beta_i(\cdot,\cdot,\cdot) > 0$. For the chosen model (29) this requirement may not hold for the values of d, k are unknown, and hence we cannot tell if $\beta_1 = d - r(x_0(t),k)$ remains positive at all t. Let us therefore transform (29) into (2) by the following parameter-dependent non-singular transformation:

$$x_1 = r_{\max}(bs_1 + s_0),\ \ x_0 = s_0.$$

In the new coordinates the system dynamics is described as follows:

$$
\begin{aligned}
\dot{x}_0 &= -dx_0 + du(t) - \frac{x_0}{x_0+k}x_1 + r_{\max}\frac{x_0}{x_0+k}x_0 \\
\dot{x}_1 &= -dx_1 + r_{\max}du(t), \\
y &= x_0,
\end{aligned}
\tag{30}
$$

Denoting $\theta_{0,1} = -d$, $\theta_{0,2} = d$, $q_1 = k$, $c_1(x_0, q_1) = -x_0/(x_0+k)$, $\tau_1 = d$, $\theta_{1,1} = r_{\max}d$, $p_0 = k$, $c_0(x_0, q_0) = r_{\max}x_0^2/(x_0+k)$ we can rewrite (30) as

$$
\begin{aligned}
&\dot{x}_0 = \theta_0^T\phi_0(x_0, p_0, t) + c_1(x_0, q_1)x_1 + c_0(x_0, q_0) \\
&\dot{x}_1 = -\tau_1 x_1 + \theta_{1,1}\phi_1(t), \ \boldsymbol{\theta} = (\theta_{0,1}, \theta_{0,2}, \theta_{1,1})^T, \ \theta_0 = (\theta_{0,1}, \theta_{0,2}),
\end{aligned}
\tag{31}
$$

where

$$\phi_0(x_0, p_0, t) = (x_0, u(t)), \ \phi_1(t) = u(t).$$

Taking (6) into account, and that $p_0 = q_0 = q_1 = k$ we define

$$
\begin{aligned}
&\bar{\boldsymbol{\varphi}}(x_0, \boldsymbol{\lambda}, t) = \left(x_0(t), u(t), -\frac{x_0(t)}{x_0(t)+p_0}\int_{t_0}^{t} e^{-\tau_1(t-\chi)}u(\chi)d\chi\right), \\
&\qquad \boldsymbol{\lambda} = (p_0, \tau_1) = (\lambda_1, \lambda_2).
\end{aligned}
$$

According to (11) – (13), the structure of the observer is:

$$
\begin{cases}
\dot{\hat{x}}_0 = -(\hat{x}_0 - x_0) + \sum_{i=1}^{3}\hat{\theta}_i\bar{\varphi}_i(x_0, \boldsymbol{\lambda}, t) + c_0(x_0, \boldsymbol{\lambda}) \\
\dot{\hat{\theta}}_i = -\gamma_\theta(\hat{x}_0 - x_0)\bar{\varphi}_i(x_0, \boldsymbol{\lambda}, t)
\end{cases}
\tag{32}
$$

$$
\begin{cases}
\dot{\hat{x}}_{1,1} &= \gamma \cdot \omega_1 \cdot e \cdot \left(\hat{x}_{1,1} - \hat{x}_{2,1} - \hat{x}_{1,1}\left(\hat{x}_{1,1}^2 + \hat{x}_{2,1}^2\right)\right) \\
\dot{\hat{x}}_{2,1} &= \gamma \cdot \omega_1 \cdot e \cdot \left(\hat{x}_{1,1} + \hat{x}_{2,1} - \hat{x}_{2,1}\left(\hat{x}_{1,1}^2 + \hat{x}_{2,1}^2\right)\right) \\
\dot{\hat{x}}_{1,2} &= \gamma \cdot \omega_2 \cdot e \cdot \left(\hat{x}_{1,2} - \hat{x}_{2,2} - \hat{x}_{1,2}\left(\hat{x}_{1,2}^2 + \hat{x}_{2,2}^2\right)\right) \\
\dot{\hat{x}}_{2,2} &= \gamma \cdot \omega_2 \cdot e \cdot \left(\hat{x}_{1,2} + \hat{x}_{2,2} - \hat{x}_{2,2}\left(\hat{x}_{1,2}^2 + \hat{x}_{2,2}^2\right)\right) \\
\hat{\lambda}_1(\hat{x}_{1,1}) &= \lambda_{\min,1} + \frac{\lambda_{\max,1} - \lambda_{\min,1}}{2}(\hat{x}_{1,1} + 1), \\
\hat{\lambda}_2(\hat{x}_{1,2}) &= \lambda_{\min,2} + \frac{\lambda_{\max,2} - \lambda_{\min,2}}{2}(\hat{x}_{1,2} + 1), \\
e &= \tanh(\|x_0(t) - \hat{x}_0(t)\|_\varepsilon),
\end{cases}
\tag{33}
$$

Parameters ω_1, ω_2 in (33) are to be rationally independent. In our example we set these parameters as follows: $\omega_1 = \pi$, $\omega_2 = 2\sqrt{2}$. Other parameters of the observer were set as $\gamma = 0.0001$, $\epsilon = 0.03$, $\gamma_\theta = 0.001$, $\alpha = 1$. Parameters d and k of the reactor model were set to 0.3 and 70 respectively, and the input $u(t) = 40(\sin(0.2t) + 1.5)$. Results of computer simulations of the closed-loop system with this observer are provided in Fig. 2. As we can see from these

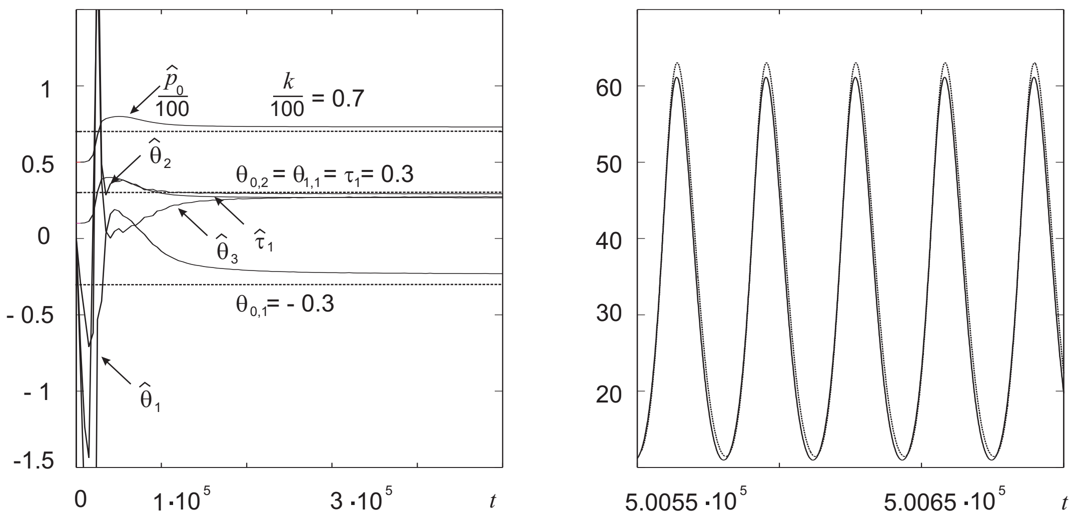

Fig. 2. Left: estimates $\hat{\theta}_i$, $\hat{\tau} = \hat{\lambda}_1$, $\hat{k} = \hat{\lambda}_2/100$ as functions of time. Right: $\hat{s}_1 = x_1 - s_0$ (dotted line) and s_1 (solid line) as functions of time in the closed-loop system (31), (32), (33).

figures, state and parameter estimates converge into a vicinity of the true values of s_1 and $\theta_1, \theta_2, \theta_3, \tau_1, k$ as $t \to \infty$.

Observer (33), with minor modifications, can be easily tailored to deal with other similar models of chemical oscillators, such as e.g. [5], in which the function $r(s_0, k)$ is replaced with an exponential nonlinearity. Furthermore, it can be used to estimate state and parameter values of various models of interacting species. We illustrate this below using the Lotka-Volterra system as an example.

Example 3. Lotka-Volterra system. Consider the following system:

$$\begin{aligned} \dot{x} &= x(\alpha - y) \\ \dot{y} &= -y(\gamma - \delta x), \end{aligned} \tag{34}$$

where x, y are the variables modeling the number of prey and predators in an isolated population. Parameters α, γ, and δ are supposed to be unknown, and x is measured. We wish to be able to estimate the values of all model parameters and also recover the values of y from the measurement of x over time.

Similarly to the previous example, we introduce the following nonsingular coordinate transformation:

$$x_0 = x, \;\; x_1 = y + \delta x.$$

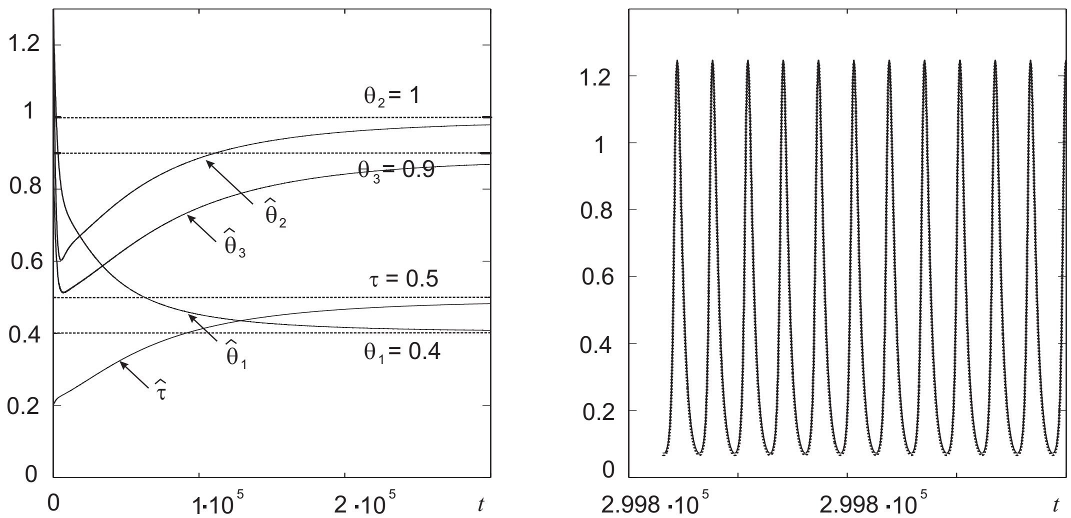

Fig. 3. Left: estimates $\hat{\theta}_i$, $\hat{\tau} = \hat{\lambda}$ as functions of time. Right: $\hat{y} = x_1 - \hat{\theta}_2 x_0$ (dotted line) and y (solid line) as functions of time in the closed-loop system (35), (36), (37).

In the new coordinates system (34) is defined as

$$\begin{aligned} &\dot{x}_0 = \theta_1 x_0 + \theta_2 x_0^2 + c_1(x_0) x_1 \\ &\dot{x}_1 = -\tau_1 x_1 + \theta_2 x_0, \\ &\theta_1 = \alpha,\ \theta_2 = \delta,\ \tau_1 = \gamma,\ \theta_3 = \delta(\alpha + \gamma),\ c_1(x_0) = -x_0. \end{aligned} \tag{35}$$

According to (11) – (13), the following system can serve as an observer for (34)

$$\begin{cases} \dot{\hat{x}}_0 = -(\hat{x}_0 - x_0) + \sum_{i=1}^{3} \hat{\theta}_i \bar{\varphi}_i(x_0, \boldsymbol{\lambda}, t) \\ \dot{\hat{\theta}}_i = -\gamma_\theta (\hat{x}_0 - x_0) \bar{\varphi}_i(x_0, \boldsymbol{\lambda}, t) \end{cases} \tag{36}$$

$$\begin{cases} \dot{\hat{x}}_{1,1} &= \gamma \cdot e \cdot \left(\hat{x}_{1,1} - \hat{x}_{2,1} - \hat{x}_{1,1}\left(\hat{x}_{1,1}^2 + \hat{x}_{2,1}^2\right)\right) \\ \dot{\hat{x}}_{2,1} &= \gamma \cdot e \cdot \left(\hat{x}_{1,1} + \hat{x}_{2,1} - \hat{x}_{2,1}\left(\hat{x}_{1,1}^2 + \hat{x}_{2,1}^2\right)\right) \\ \hat{\lambda}(\hat{x}_{1,1}) &= \lambda_{\min} + \frac{\lambda_{\max} - \lambda_{\min}}{2}(\hat{x}_{1,1} + 1), \\ e &= \tanh(\|x_0(t) - \hat{x}_0(t)\|_\varepsilon), \end{cases} \tag{37}$$

where

$$\bar{\boldsymbol{\varphi}}(x_0(t), \boldsymbol{\lambda}, t) = \left(x_0(t), x_0^2(t), -x_0(t) \int_{t_0}^{t} e^{-\tau_1(t-\chi)} x_0(\chi) d\chi\right)$$

We simulated system (35), (36), (37) with the following parameters of the observer: $\epsilon = 0.0005$, $\gamma = 0.0015$, $\gamma_\theta = 0.02$, $\alpha = 1$, $\lambda_{\min} = 0.2$, $\lambda_{\max} = 0.6$. Results of the simulation are provided in Fig. 3.

6 Conclusion

We derived, under suitable assumptions, an observer that can reconstruct the unknown parameters of a general class of nonlinear systems, specified by equation (2). Reconstruction is subject to conditions of linear/nonlinear persistency of excitation. According to Theorem 5, the values of parameters can be reconstructed with arbitrarily small errors (up to their equivalence class). Hence there exists a time instance t' such that solutions of (12) will enter a sufficiently small neighborhood of $x_i(t)$ at $t = t'$ and will stay there for all $t \geq t'$ provided that the values of Δ_φ, Δ_ξ in (8), (10) are sufficiently small. Having solved the parameter reconstruction problem, we obtain the state variables by applying the filter (12).

To ensure convergence we have gone beyond the usual requirement of Lyapunov stability. The set to which the estimates converge is not stable in the Lyapunov sense; any neighborhood of this set contains a nonempty subset of points corresponding to trajectories escaping this neighborhood in finite time. Yet the set is attracting in a weaker, Milnor sense, cf. [22], and furthermore it is made Poisson stable by design. The amount of time required for convergence in our approach depends heavily on the subspace being searched, it does not depend on the dimension of the linearly parameterized part. This renders the method more efficient than exhaustive search; the fewer uncertainties, relatively speaking, belong to the nonlinear part, the more advantageous our method becomes.

Acknowledgment

The authors are grateful to Prof. A.N. Gorban, the University of Leicester, and V.A. Terekhov, Staint-Petersburg State University of Electrical Engineering, for their useful comments and numerous technical discussions during preparation of this work. The research was supported by a Royal Society International Joint Project grant, and partially supported by a grant from Royal Academy of Engineering and RFBR grant 8-08-00103-a.

Appendix

A.1 Proof of Theorem 3

In order to prove the theorem we will need the following lemma.

Lemma 6 *Let the system (11),(13) be given, and the function* $\bar{\boldsymbol{\varphi}}(x_0(t), \boldsymbol{\lambda}, t)$ *be* λ*-uniformly persistently exciting. Then there exist numbers* $\rho > 0$*,* $D_\rho > 0$*,* $c_1 > 0$*,* $c_2 > 0$*,* $\bar{\gamma} > 0$*, such that for all* $\gamma_i \in (0, \bar{\gamma})$ *the following holds along the solutions of system (2), (11), (13):*

$$
\begin{aligned}
\|x_0(t) - \hat{x}_0(t)\| &\leq e^{-\rho(t-t_0)} D_\rho(\|x(t_0) - \hat{x}(t_0)\| + \|\hat{\boldsymbol{\theta}}(t) - \hat{\boldsymbol{\theta}}(t_0)\|) \\
&\quad + c_1 \|\hat{\boldsymbol{\lambda}}(\tau) - \boldsymbol{\lambda}\|_{\infty,[t_0,t]} + d_1 \\
\|\hat{\boldsymbol{\theta}}(t) - \hat{\boldsymbol{\theta}}(t_0)\| &\leq e^{-\rho(t-t_0)} D_\rho(\|x(t_0) - \hat{x}(t_0)\| + \|\hat{\boldsymbol{\theta}}(t) - \hat{\boldsymbol{\theta}}(t_0)\|) \\
&\quad + c_2 \|\hat{\boldsymbol{\lambda}}(\tau) - \boldsymbol{\lambda}\|_{\infty,[t_0,t]} + d_2
\end{aligned}
\tag{A.1}
$$

$$
c_1 = c_2 = (\|\boldsymbol{\theta}\| D + D_c) D_\rho \rho^{-1}, \quad d_1 = d_2 = 2(\|\boldsymbol{\theta}\| \Delta_\varphi + \Delta_\xi) D_\rho \rho^{-1}
$$

Proof of Lemma 6. Consider the following functions

$$
\begin{aligned}
\boldsymbol{\eta}(\boldsymbol{\lambda}, t) &= \bar{\boldsymbol{\varphi}}(x_0(t), \boldsymbol{\lambda}, t) - \boldsymbol{\varphi}(x_0(t), \boldsymbol{\lambda}, t) \\
\chi(t) &= \sum_{i=1}^{n} c_i(x_0(t), q_i, t) \int_{t_0}^{t} e^{-\int_\tau^t \beta_i(x_0(s), \tau_i, s) ds} \xi_i(\tau) d\tau + \xi_0(t)
\end{aligned}
\tag{A.2}
$$

According to (8), (10) the functions $\boldsymbol{\eta}(\boldsymbol{\lambda}, t)$, $\chi(t)$ are bounded: $\|\boldsymbol{\eta}(\boldsymbol{\lambda}, \tau)\|_{\infty,[0,t]} \leq \Delta_\varphi$, $\|\chi(\tau)\|_{\infty,[0,t]} \leq \Delta_\xi$. Now introduce the following vector

$$
\mathbf{q} = \mathrm{col}(\hat{x}_0 - x_0, \hat{\boldsymbol{\theta}}^T - \boldsymbol{\theta}^T).
\tag{A.3}
$$

Taking into account (2), (11), its time-derivative can be expressed as follows:

$$
\begin{aligned}
\dot{q}_1 &= -\alpha\ q_1 + \hat{\boldsymbol{\theta}}^T \bar{\boldsymbol{\varphi}}(x_0, \hat{\boldsymbol{\lambda}}, t) + c_0(x_0(t), \hat{\boldsymbol{\lambda}}, t) - \theta_0^T \boldsymbol{\phi}_0(x_0(t), p_0, t) \\
&\quad - \sum_{j=1}^{n} c_j(x_0(t), q_j, t) x_j(t) - c_0(x_0, \boldsymbol{\lambda}, t) - \xi_0(t) \\
\dot{q}_i &= -\gamma_\theta\ q_1\ \bar{\boldsymbol{\varphi}}_i(x_0(t), \hat{\boldsymbol{\lambda}}, t),\ i = \{1, \ldots, n\}
\end{aligned}
\tag{A.4}
$$

Expressing trajectories $x_i(t)$ in (2) in the closed form

$$
\begin{aligned}
x_i(t) &= e^{-\int_{t_0}^t \beta_i(x_0(s), \tau_i, s) ds} x_i(t_0) + c_i(x_0, q_i, t) \int_{t_0}^{t} e^{-\int_\tau^t \beta_i(x_0(s), \tau_i, s) ds} \xi_i(\tau) d\tau + \\
&\quad \theta_i^T c_i(x_0, q_i, t) \int_{t_0}^{t} e^{-\int_\tau^t \beta_i(x_0(s), \tau_i, s) ds} \boldsymbol{\phi}_i(x_0(\tau), p_i, \tau) d\tau,
\end{aligned}
$$

and taking (7), (A.2) into account, we can rewrite (A.4) as

$$
\begin{aligned}
\dot{q}_1 &= -\alpha\ q_1 + (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^T \bar{\boldsymbol{\varphi}}(x_0, \hat{\boldsymbol{\lambda}}, t) + \upsilon(\hat{\boldsymbol{\lambda}}, \boldsymbol{\lambda}, t) \\
\dot{q}_i &= -\gamma_\theta\ q_1\ \bar{\boldsymbol{\varphi}}_i(x_0(t), \hat{\boldsymbol{\lambda}}, t),\ i = \{1, \ldots, n\},
\end{aligned}
\tag{A.5}
$$

where

$$
\begin{aligned}
\upsilon(\hat{\boldsymbol{\lambda}}, \boldsymbol{\lambda}, t) &= \boldsymbol{\theta}^T(\bar{\boldsymbol{\varphi}}(x_0, \hat{\boldsymbol{\lambda}}, t) - \bar{\boldsymbol{\varphi}}(x_0, \boldsymbol{\lambda}, t)) + c_0(x_0, \boldsymbol{\lambda}, t) - c_0(x_0, \hat{\boldsymbol{\lambda}}, t) \\
&\quad + \boldsymbol{\theta}^T \boldsymbol{\eta}(t, \boldsymbol{\lambda}) - \chi(t) + \epsilon(t)
\end{aligned}
\tag{A.6}
$$

and $\epsilon(t)$ is a bounded and exponentially decaying term. Rewriting (A.5) in vector-matrix notation yields:

$$\dot{\mathbf{q}} = \mathbf{A}(\hat{\boldsymbol{\lambda}}(t), t)\mathbf{q} + \mathbf{b}\ \upsilon(\boldsymbol{\lambda}, \hat{\boldsymbol{\lambda}}, t), \tag{A.7}$$

where

$$\begin{aligned} \mathbf{A}(\hat{\boldsymbol{\lambda}}(t), t) &= \begin{pmatrix} -\alpha & \bar{\boldsymbol{\varphi}}(x_0(t), \hat{\boldsymbol{\lambda}}(t), t)^T \\ -\gamma_\theta \bar{\boldsymbol{\varphi}}(x_0(t), \hat{\boldsymbol{\lambda}}(t), t) & 0 \end{pmatrix}, \\ \mathbf{b} &= (1, 0, \ldots, 0)^T. \end{aligned} \tag{A.8}$$

Let $\Phi(t, t_0)$ be a solution of

$$\dot{\Phi} = \mathbf{A}(\hat{\boldsymbol{\lambda}}(t), t)\Phi,\ \ \Phi(t_0, t_0) = \mathbf{I} \tag{A.9}$$

Then the solution of (A.7) is defined as

$$\begin{aligned} \mathbf{q}(t) &= \Phi(t, t_0)\mathbf{q}(t_0) + \int_{t_0}^{t} \Phi(t, \tau)\mathbf{b}\upsilon(\hat{\boldsymbol{\lambda}}(\tau), \boldsymbol{\lambda}, \tau) d\tau, \\ t &\geq t_0 \end{aligned} \tag{A.10}$$

We are going to show that there exists $\bar{\gamma} \in \mathbb{R}_{>0}$ such that for all $\gamma \in [0, \bar{\gamma})$ solutions of (A.7), (13) are bounded.

First, we notice that trajectories $\hat{x}_{1,j}(t)$, $\hat{x}_{2,j}(t)$ are globally bounded. Furthermore, the right-hand side of (13) is locally Lipschitz in $\hat{x}_{1,j}$, $\hat{x}_{2,j}$. Hence the following estimate holds:

$$\|\dot{\hat{\boldsymbol{\lambda}}}(t)\| \leq \gamma^* M,\ \ M \in \mathbb{R}_{>0}, \gamma^* = \gamma \max_j \{\omega_j\}. \tag{A.11}$$

As follows from the assumptions of the lemma, the function $\bar{\boldsymbol{\varphi}}(x_0(t), \boldsymbol{\lambda}, t)$ is a λ-uniform PE. This implies existence of $L, \mu \in \mathbb{R}_{>0}$ such that

$$\begin{aligned} J(\boldsymbol{\lambda}, t) = \int_{t}^{t+L} & \bar{\boldsymbol{\varphi}}(x_0(\tau), \boldsymbol{\lambda}, \tau)\bar{\boldsymbol{\varphi}}^T(x_0(\tau), \boldsymbol{\lambda}, \tau) d\tau \geq \mu I \\ & \forall\ t \geq t_0,\ \ \boldsymbol{\lambda} \in \Omega_\lambda \end{aligned} \tag{A.12}$$

Consider the following matrix:

$$
\begin{aligned}
&J(\hat{\boldsymbol{\lambda}}(t),t) - \int_t^{t+L} \bar{\boldsymbol{\varphi}}(x_0(\tau),\hat{\boldsymbol{\lambda}}(\tau),\tau)\bar{\boldsymbol{\varphi}}^T(x_0(\tau),\hat{\boldsymbol{\lambda}}(\tau),\tau)d\tau \\
&= \int_t^{t+L} (\bar{\boldsymbol{\varphi}}(x_0(\tau),\hat{\boldsymbol{\lambda}}(t),\tau) - \bar{\boldsymbol{\varphi}}(x_0(\tau),\hat{\boldsymbol{\lambda}}(\tau),\tau)) \times \\
&\bar{\boldsymbol{\varphi}}^T(x_0(\tau),\hat{\boldsymbol{\lambda}}(t),\tau)d\tau + \int_t^{t+L} \bar{\boldsymbol{\varphi}}(x_0(\tau),\hat{\boldsymbol{\lambda}}(\tau),\tau) \times \\
&(\bar{\boldsymbol{\varphi}}(x_0(\tau),\hat{\boldsymbol{\lambda}}(t),\tau) - \bar{\boldsymbol{\varphi}}(x_0(\tau),\hat{\boldsymbol{\lambda}}(\tau),\tau))^T d\tau \\
&= J_1(\hat{\boldsymbol{\lambda}}(t),t) + J_2(\hat{\boldsymbol{\lambda}}(t),t)
\end{aligned} \tag{A.13}
$$

Using the inequality

$$
\|H\mathbf{z}\| \leq \max_{k,l} |h_{k,l}| \|\mathbf{z}\|, \ H \in \mathbb{R}^{m\times n}, \ \mathbf{z} \in \mathbb{R}^n
$$

and given that $\|\bar{\boldsymbol{\varphi}}(x_0(t),\boldsymbol{\lambda},t)\| \leq B$ for all $t \geq 0$, $\boldsymbol{\lambda} \in \Omega_\lambda$, we can conclude that the matrix (A.13) satisfies

$$
\begin{aligned}
&\left|\mathbf{z}^T(J_1(\boldsymbol{\lambda}(t),t) + J_2(\boldsymbol{\lambda}(t),t))\mathbf{z}\right| \leq \\
&\leq 2BD\|\hat{\boldsymbol{\lambda}}(t) - \hat{\boldsymbol{\lambda}}(\tau)\|_{\infty,[t,t+L]}\|\mathbf{z}\|^2, \ D \in \mathbb{R}_{>0}.
\end{aligned}
$$

Hence

$$
\begin{aligned}
&\int_t^{t+L} \bar{\boldsymbol{\varphi}}(x_0(\tau),\hat{\boldsymbol{\lambda}}(\tau),\tau)\bar{\boldsymbol{\varphi}}^T(x_0(\tau),\hat{\boldsymbol{\lambda}}(\tau),\tau)d\tau \\
&\geq J(\hat{\boldsymbol{\lambda}}(t),t) - 2BD\|\hat{\boldsymbol{\lambda}}(t) - \hat{\boldsymbol{\lambda}}(\tau)\|_{\infty,[t,t+L]} I \geq \\
&(\mu - 2BD\|\hat{\boldsymbol{\lambda}}(t) - \hat{\boldsymbol{\lambda}}(\tau)\|_{\infty,[t,t+L]}) I
\end{aligned} \tag{A.14}
$$

Taking (A.11), (A.14) into account we can conclude that

$$
\begin{aligned}
&\int_t^{t+L} \bar{\boldsymbol{\varphi}}(x_0(\tau),\hat{\boldsymbol{\lambda}}(\tau),\tau)\bar{\boldsymbol{\varphi}}^T(x_0(\tau),\hat{\boldsymbol{\lambda}}(\tau),\tau)d\tau \\
&\geq (\mu - 2BDLM\gamma^*) I
\end{aligned} \tag{A.15}
$$

Then choosing γ for instance such that

$$
\gamma^* = \frac{\mu}{4BDLM} \tag{A.16}
$$

we can ensure that

$$
\int_t^{t+L} \bar{\boldsymbol{\varphi}}(x_0(\tau),\hat{\boldsymbol{\lambda}}(\tau),\tau)\bar{\boldsymbol{\varphi}}^T(x_0(\tau),\hat{\boldsymbol{\lambda}}(\tau),\tau)d\tau \geq \frac{\mu}{2} I \ \forall \ t \geq t_0.
$$

In other words the function $\bar{\boldsymbol{\varphi}}(x_0(t),\hat{\boldsymbol{\lambda}}(t),t)$ is persistently exciting.

According to [23], this implies that the system

$$
\dot{\mathbf{q}} = \mathbf{A}(\hat{\boldsymbol{\lambda}}(t),t)\mathbf{q}
$$

is uniformly exponentially stable for all $\hat{\boldsymbol{\lambda}}(t)$ satisfying conditions (A.11), (A.16). Hence there exist $\rho,\ D_\rho \in \mathbb{R}_{>0}$ such that the following inequality holds

$$\|\Phi(t,t_0)\mathbf{q}(t_0)\| \leq e^{-\rho(t-t_0)}\|\mathbf{q}(t_0)\|D_\rho \tag{A.17}$$

for all $\mathbf{q}(t_0) \in \mathbb{R}^n$, $t \geq t_0$, $t_0 \in \mathbb{R}$. Therefore, taking (A.10), (A.17) into account we can conclude that, for any bounded and continuous function $\upsilon(t)$, solutions of (A.7) satisfy the following estimate:

$$\begin{aligned}
&\|\mathbf{q}(t)\| = \left\|\Phi(t,t_0)\mathbf{q}(t_0) + \int_{t_0}^{t}\Phi(t,\tau)\mathbf{b}\upsilon(\tau)d\tau\right\| \\
&\leq e^{-\rho(t-t_0)}D_\rho\|\mathbf{q}(t_0)\| + D_\rho\int_{t_0}^{t}e^{-\rho(t-\tau)}\|\mathbf{b}\||\upsilon(\tau)|d\tau.
\end{aligned} \tag{A.18}$$

Notice that according to (10), (19), (A.6) the function $\upsilon(t)$ satisfies the following inequality:

$$|\upsilon(t)| \leq (\|\boldsymbol{\theta}\|D + D_c)\|\hat{\boldsymbol{\lambda}}(t) - \boldsymbol{\lambda}\| + \|\boldsymbol{\theta}\|\Delta_\varphi + \Delta_\xi + |\epsilon(t)|.$$

Hence

$$\begin{aligned}
\|\mathbf{q}(t)\| &\leq e^{-\rho(t-t_0)}D_\rho\|\mathbf{q}(t_0)\| + (\|\boldsymbol{\theta}\|D + D_c)D_\rho\int_{t_0}^{t}e^{-\rho(t-\tau)}\|\hat{\boldsymbol{\lambda}}(\tau) - \boldsymbol{\lambda}\|d\tau \\
&+ \int_{t_0}^{t}e^{-\rho(t-\tau)}\left((\|\boldsymbol{\theta}\|\Delta_\varphi + \Delta_\xi)D_\rho + \epsilon(\tau)\right)d\tau.
\end{aligned}$$

Because $\epsilon(t)$ is exponentially decaying, there exists t' such that

$$|\epsilon(t)| \leq (\|\boldsymbol{\theta}\|\Delta_\varphi + \Delta_\xi)D_\rho$$

for all $t \geq t'$. Therefore for all $t \geq t_0 \geq t'$ the following estimate holds:

$$\begin{aligned}
\|\mathbf{q}(t)\| &\leq e^{-\rho(t-t_0)}D_\rho\|\mathbf{q}(t_0)\| + \frac{(\|\boldsymbol{\theta}\|D + D_c)D_\rho}{\rho}\|\hat{\boldsymbol{\lambda}}(\tau) - \boldsymbol{\lambda}\|_{\infty,[t_0,t]} + \\
&\frac{2(\|\boldsymbol{\theta}\|\Delta_\varphi + \Delta_\xi)D_\rho}{\rho}.
\end{aligned} \tag{A.19}$$

Noticing that $\|\mathbf{q}(t_0)\| = \sqrt{(x_0(t_0) - \hat{x}_0(t_0))^2 + \|\boldsymbol{\theta}(t_0) - \boldsymbol{\theta}\|^2}$ and using the inequality $\sqrt{a^2+b^2} \leq \sqrt{a^2} + \sqrt{b^2}$ we can conclude from (A.19) that there exists $t' \in \mathbb{R}$ such that for all $t \geq t_0 \geq t'$ estimate (A.1) holds $\square$

According to Lemma 6 there exists a non-empty interval $(0,\bar{\gamma})$ and $t' \in \mathbb{R}$ such that for all $\gamma \in (0,\bar{\gamma})$ and all $t \geq t_0 \geq t'$ solutions of system (2), (11), (13) satisfy inequality (A.1).

Now consider solutions of system

$$\begin{aligned} \dot{\bar{x}}_{1,j} &= \omega_j \left(\bar{x}_{1,j} - \bar{x}_{2,j} - \bar{x}_{1,j} \left(\hat{x}_{1,j}^2 + \bar{x}_{2,j}^2 \right) \right) \\ \dot{\bar{x}}_{2,j} &= \omega_j \left(\bar{x}_{1,j} + \bar{x}_{2,j} - \bar{x}_{2,j} \left(\bar{x}_{1,j}^2 + \bar{x}_{2,j}^2 \right) \right) \\ \bar{\lambda}_j(\bar{x}_{1,j}) &= \lambda_{j,\min} + \frac{\lambda_{j,\max} - \lambda_{j,\min}}{2} (\bar{x}_{1,j} + 1) \end{aligned} \tag{A.20}$$

with initial conditions (14). They are forward-invariant on $\bar{x}_{1,j}^2(t) + \bar{x}_{2,j}^2(t) = 1$ and can be expressed as $\bar{x}_{1,j}(t) = \sin(\omega_j t + \beta)$, $\bar{x}_{2,j}(t) = \cos(\omega_j t + \beta)$. Taking into account that ω_j are rationally-independent (condition (15) in the definition of the observer) we can conclude that the trajectories $\bar{x}_{1,j}(t)$ densely fill an invariant n-dimensional torus [1], and system (A.20) with initial conditions $\bar{x}_{1,j}^2(t_0) + \bar{x}_{2,j}^2(t_0) = 1$ is Poisson-stable in $\Omega_x = \{\bar{x}_{1,j}, \bar{x}_{2,j} \in \mathbb{R} | \bar{x}_{1,j} \in [-1, 1]\}$. This means that for any $\boldsymbol{\lambda} \in \Omega_\lambda$, (arbitrarily large) constant $T \in \mathbb{R}_{\geq 0}$, (arbitrarily small) constant $\Delta_\lambda \in \mathbb{R}_{\geq 0}$, and initial conditions on the torus there will exist $\bar{\boldsymbol{\lambda}}(t') = \mathrm{col}(\bar{\lambda}_1(\bar{x}_{1,1}(t')), \ldots, \bar{\lambda}_s(\bar{x}_{1,s}(t'))) \in \Omega_\lambda$, $t' - t_0 \geq T$, such that $\|\boldsymbol{\lambda} - \bar{\boldsymbol{\lambda}}(t')\| \leq \Delta_\lambda$, $\Delta_\lambda \in \mathbb{R}_{\geq 0}$. Expressing $\hat{\boldsymbol{\lambda}}(t)$ generated by (13) in terms of the function $\bar{\boldsymbol{\lambda}}(t)$ yields:

$$\hat{\boldsymbol{\lambda}}(t) = \bar{\boldsymbol{\lambda}} \left(t_0 + \gamma \int_{t_0}^{t} \|x_0(\tau) - \hat{x}_0(\tau)\|_\varepsilon d\tau \right), \ t \geq t_0.$$

Denoting

$$h(t) = t' - \gamma \int_{t_0}^{t} \|x_0(\tau) - \hat{x}_0(\tau)\|_\varepsilon d\tau$$

we obtain

$$\begin{aligned} \|\hat{\boldsymbol{\lambda}}(t) - \boldsymbol{\lambda}\| &\leq \Delta_\lambda + \|\hat{\boldsymbol{\lambda}}(t) - \bar{\boldsymbol{\lambda}}(t')\| \leq \Delta_\lambda + D_\lambda |h(t)| \\ h(t) &= h(t_0) - \gamma \int_{t_0}^{t} \|x_0(\tau) - \hat{x}_0(\tau)\|_\varepsilon d\tau, \end{aligned} \tag{A.21}$$

where D_λ is a positive constant and $h(t_0) = t'$. Hence the following holds along the solutions of (2), (11), (13):

$$\begin{aligned} \|x_0(t) - \hat{x}_0(t)\| &\leq e^{-\rho(t-t_0)} D_\rho (\|x(t_0) - \hat{x}(t_0)\| + \|\hat{\boldsymbol{\theta}}(t) - \hat{\boldsymbol{\theta}}(t_0)\|) \\ &\quad + c_1 D_\lambda \|h(t)\|_{\infty,[t_0,t]} + d_1 + c_1 \Delta_\lambda \\ \|\hat{\boldsymbol{\theta}}(t) - \hat{\boldsymbol{\theta}}(t_0)\| &\leq e^{-\rho(t-t_0)} D_\rho (\|x(t_0) - \hat{x}(t_0)\| + \|\hat{\boldsymbol{\theta}}(t) - \hat{\boldsymbol{\theta}}(t_0)\|) \\ &\quad + c_2 D_\lambda \|h(t)\|_{\infty,[t_0,t]} + d_2 + c_2 \Delta_\lambda \end{aligned} \tag{A.22}$$

where $h(t)$ is defined as in (A.21), $c_1 = c_2 = (\|\boldsymbol{\theta}\| D + D_c) D_\rho \rho^{-1}$, and Δ_λ is an arbitrarily small constant. Therefore, according to Lemma 8 in Section A.3 of the Appendix and invoking the argument as in Corollary 11 from [28], we can conclude that there exists γ^* and ε^* such that

$$\lim_{t \to \infty} \hat{\boldsymbol{\lambda}}(t) = \boldsymbol{\lambda}^*, \ \boldsymbol{\lambda}^* \in \Omega_\lambda \tag{A.23}$$

$$\lim_{t\to\infty} \|x_0(t) - \hat{x}_0(t)\|_\varepsilon = 0. \tag{A.24}$$

for all $\gamma \in (0, \gamma^*)$, $\varepsilon \geq \varepsilon^*$. Then denoting $\kappa = D_\rho/\rho$ and using (A.19) we obtain (21). The value of γ^*, as follows from Lemma 8 and (A.16) can be determined from

$$\gamma^* = \min\left\{\frac{\mu}{4BDLM}, \mathcal{G}\right\}$$
$$\mathcal{G} = \inf_{d\in(0,1),\ \psi\in(1,\infty)} \frac{\psi-1}{\psi}\rho^2 \left[\ln\left(\frac{2\psi D_\rho}{d}\right)\right]^{-1} \frac{1}{2(\|\boldsymbol{\theta}\|D + D_c)D_\rho D_\lambda}\left(1 + \frac{D_\rho \psi}{1-d}\right)^{-1}$$

$\square$

A.2 Proof of Theorem 5

Similar to the proof of Theorem 3, equation (A.7), we notice that equations governing the dynamics of our observer (11), (13) relative to the dynamics of the original system can be described as follows

$$\begin{aligned} \dot{\mathbf{q}} &= \mathbf{A}(\hat{\boldsymbol{\lambda}}(t), t)\mathbf{q} + \mathbf{b}\upsilon(\hat{\boldsymbol{\lambda}}, \boldsymbol{\lambda}, t) \\ y &= \mathbf{c}^T\mathbf{q}, \end{aligned} \tag{A.25}$$

where

$$\mathbf{c} = \mathrm{col}(1, 0, \ldots, 0),\ \mathbf{b} = \mathrm{col}(1, 0, \ldots, 0),\ \mathbf{q} = \mathrm{col}(\hat{x}_0 - x_0, \hat{\boldsymbol{\theta}}^T - \boldsymbol{\theta}^T),$$

and system

$$\dot{\mathbf{q}} = \mathbf{A}(\hat{\boldsymbol{\lambda}}(t), t)\mathbf{q}$$

is uniformly exponentially stable. Taking (A.6) into account we can express $\upsilon(\hat{\boldsymbol{\lambda}}, \boldsymbol{\lambda}, t)$ as:

$$\begin{aligned} \upsilon(\hat{\boldsymbol{\lambda}}, \boldsymbol{\lambda}, t) = \boldsymbol{\theta}^T(\bar{\boldsymbol{\varphi}}(x_0, \hat{\boldsymbol{\lambda}}, t) - \bar{\boldsymbol{\varphi}}(x_0, \boldsymbol{\lambda}, t)) + c_0(x_0, \hat{\boldsymbol{\lambda}}, t) - c_0(x_0, \boldsymbol{\lambda}, t) + \\ \boldsymbol{\theta}^T\boldsymbol{\eta}(t, \boldsymbol{\lambda}) - \chi(t) + \epsilon(t) = \upsilon_0(\hat{\boldsymbol{\lambda}}, \boldsymbol{\lambda}, t) + d(\boldsymbol{\theta}, \boldsymbol{\lambda}, t) \end{aligned}$$

In the equation above $\upsilon_0(\hat{\boldsymbol{\lambda}}, \boldsymbol{\lambda}, t)$ denotes the term $\boldsymbol{\theta}^T(\bar{\boldsymbol{\varphi}}(x_0, \hat{\boldsymbol{\lambda}}, t) - \bar{\boldsymbol{\varphi}}(x_0, \boldsymbol{\lambda}, t)) + c_0(x_0, \hat{\boldsymbol{\lambda}}, t) - c_0(x_0, \boldsymbol{\lambda}, t)$, and $d(\boldsymbol{\theta}, \boldsymbol{\lambda}, t)$ stands for $\boldsymbol{\theta}^T\boldsymbol{\eta}(t, \boldsymbol{\lambda}) - \chi(t) + \epsilon(t)$. Hence (A.25) transforms into

$$\begin{aligned} \dot{\mathbf{q}} &= \mathbf{A}(\hat{\boldsymbol{\lambda}}(t), t)\mathbf{q} + \mathbf{b}[\upsilon_0(\hat{\boldsymbol{\lambda}}, \boldsymbol{\lambda}, t) + d(\boldsymbol{\theta}, \boldsymbol{\lambda}, t)] \\ y &= \mathbf{c}^T\mathbf{q}. \end{aligned} \tag{A.26}$$

To proceed further we will need the following lemma

Lemma 7 *Consider the following system*

$$\begin{cases} \dot{\mathbf{x}} = A(t)\mathbf{x} + \mathbf{b}(t)[u(t) + d(t)], \\ y = \mathbf{c}^T\mathbf{x}, \quad \mathbf{c} \in \mathbb{R}^n, \ \mathbf{x} \in \mathbb{R}^n \end{cases} \tag{A.27}$$

where

$$\begin{aligned} u :& \mathbb{R} \to \mathbb{R}, \ u \in \mathcal{C}^1, \ \mathbf{b} : \mathbb{R} \to \mathbb{R}^n, \ A(t) : \mathbb{R} \to \mathbb{R}^{n\times n} \\ d :& \mathbb{R} \to \mathbb{R}, \ \mathbf{d} \in \mathcal{C}^0, \ \|d(t)\|_{\infty,[t_0,\infty]} \leq \delta, \ \delta \in \mathbb{R}_{\geq 0} \end{aligned} \tag{A.28}$$

and each entry of $A(t)$ is a globally bounded function of t. Suppose that the following properties hold for (A.27):

1) *the origin of system (A.27) is asymptotically stable at $u(t) = 0$, $d(t) = 0$*
2) *the term $\mathbf{c}^T\mathbf{b}(t)$ is separated away from zero:*

$$\exists \ \beta \in \mathbb{R}_{>0} : \ |\mathbf{c}^T\mathbf{b}(t)| \geq \beta \ \forall \ t \in \mathbb{R}$$

3) *the time derivative of $u(t)$ is bounded uniformly in t:*

$$\exists \ \partial U \in \mathbb{R}_{>0} : \quad |\dot{u}(t)| \leq \partial U \ \ \forall \ t \in \mathbb{R} \tag{A.29}$$

Then

$$\|y(t)\|_{\infty,[t_0,\infty]} < \varepsilon, \ \ \varepsilon \in \mathbb{R}_{>0} \tag{A.30}$$

implies existence of $t'(\varepsilon) \geq t_0$ such that

$$2\sqrt{\frac{6\varepsilon\partial U}{\beta}} + \delta \geq |u(t)|, \ \ \forall \ t \geq t'(\varepsilon)$$

provided that ε is sufficiently small.

Proof. First, notice that closed form solutions of (A.27) are

$$\mathbf{x}(t) = \Phi(t, t_0)\mathbf{x}_0 + \int_{t_0}^{t} \Phi(t, \tau)\mathbf{b}(\tau)[u(\tau) + d(\tau)]d\tau, \tag{A.31}$$

where $\Phi(t, t_0)$ is the fundamental system of solutions of (A.27) generated by

$$\dot{\Phi} = A(t)\Phi, \quad \Phi(t_0, t_0) = \Phi_0 = I. \tag{A.32}$$

According to (A.31) we have:

$$y(t) = \mathbf{c}^T\Phi(t, t_0)\mathbf{x}_0 + \int_{t_0}^{t} \mathbf{c}^T\Phi(t, \tau)\mathbf{b}(\tau)[u(\tau) + d(\tau)]d\tau$$

Consider $y(t) - y(t - T)$, $T \in \mathbb{R}_{>0}$:

$$y(t)-y(t-T) = \mathbf{c}^T \left[\Phi(t, t_0) - \Phi(t - T, t_0)\right] \mathbf{x}_0 + \int_{t-T}^{t} \mathbf{c}^T\Phi(t, \tau)\mathbf{b}(\tau)[u(\tau)+d(\tau)]d\tau \tag{A.33}$$

Denoting

$$\nu(t,T) = \mathbf{c}^T \left[\Phi(t,t_0) - \Phi(t-T,t_0)\right] \mathbf{x}_0$$

and applying the mean-value theorem to (A.33) we obtain

$$y(t) - y(t-T) = \nu(t,T) + T\mathbf{c}^T\Phi(t,\tau')\mathbf{b}(\tau')[u(\tau') + d(\tau')],\ \ \tau' \in [t, t-T]$$

Hence

$$|y(t)-y(t-T)| \geq T|\mathbf{c}^T\Phi(t,\tau')\mathbf{b}(\tau')|\ (|u(\tau')| - |d(\tau')|) - |\nu(t,T)|,\ \tau' \in [t, t-T]$$

Because (A.27) is asymptotically stable, the term $|\nu(t,T)| \to 0$ as $t \to \infty$, and there exists $t_1 \geq \mathbb{R}$ such that

$$|\nu(t,T)| < \varepsilon\ \forall\ t \geq t_1.$$

On the other hand, condition (A.30) assures the existence of t_0 such that

$$|y(t) - y(t-T)| \leq 2\varepsilon\ \forall\ t \geq t_0,\ \ \forall\ T > 0.$$

Therefore there exists $t' = \max(t_1, t_0)$ such that for all $T \geq 0$ and $t \geq t'$ the following holds:

$$3\varepsilon \geq T|\mathbf{c}^T\Phi(t,\tau')\mathbf{b}(\tau')|\ (|u(\tau')| - |d(\tau')|),\ \ \tau' \in [t-T, t]. \tag{A.34}$$

According to (A.29), (A.28), inequality (A.34) implies

$$3\varepsilon \geq T|\mathbf{c}^T\Phi(t,\tau')\mathbf{b}(\tau')|\ (|u(t)| - \delta) - T^2|\mathbf{c}^T\Phi(t,\tau')\mathbf{b}(\tau')|\ \partial U,\ \ \tau' \in [t-T, t]$$

Notice that

$$\Phi(t,t) = I\ \forall\ t \in \mathbb{R}$$

and that $\Phi(t,\tau')$ is (uniformly) continuous in τ' and $|\mathbf{c}^T\Phi(t,t)\mathbf{b}(t)| \geq \beta$. Then there exists $T' \in \mathbb{R}_{>0}$ such that

$$|\mathbf{c}^T\Phi(t,\tau')\mathbf{b}(\tau')| \geq M = \frac{\beta}{2}\ \forall\ \tau' \in [t-T', t] \tag{A.35}$$

Letting $0 < T \leq T'$ we obtain

$$\frac{3\varepsilon}{MT} + T\partial U + \delta \geq |u(t)|\ \forall\ t \geq t' \tag{A.36}$$

Optimizing the left-hand side of (A.36) for T yields

$$2\sqrt{\frac{3\varepsilon\partial U}{M}} + \delta = 2\sqrt{\frac{6\varepsilon\partial U}{\beta}} + \delta \geq |u(t)|\ \forall\ t \geq t'$$

provided that

$$\sqrt{\frac{6\varepsilon\partial U}{\beta}} \leq T', \tag{A.37}$$

i.e. when ε is sufficiently small. $\square$

As follows from Theorem 3, property (20), there exists a time instant t' such that

$$\|\mathbf{c}^T\mathbf{q}(t)\| = \|\hat{x}_0(t) - x_0(t)\| \leq 2\varepsilon \ \forall\ t \geq t'$$

In addition, vectors $\mathbf{c}$, $\mathbf{b}$ in (A.25) satisfy $\mathbf{c}^T\mathbf{b} = 1$, and $d(\boldsymbol{\theta}, \boldsymbol{\lambda}, t)$ satisfies

$$|d(\boldsymbol{\theta}, \boldsymbol{\lambda}, t)| \leq \|\boldsymbol{\theta}\|\Delta_{\varphi} + \Delta_{\xi} + |\epsilon(t)| = \Delta + |\epsilon(t)|,$$

where $\epsilon(t)$ is an exponentially decaying term and $\Delta = \|\boldsymbol{\theta}\|\Delta_{\varphi} + \Delta_{\xi}$. Let t'': $|\epsilon(t)| \leq \|\boldsymbol{\theta}\|\Delta$ for all $t \geq t''$, and $t_0 = \min\{t', t''\}$. This implies that

$$\|\mathbf{c}^T\mathbf{q}(t)\| \leq 2\varepsilon, \quad |d(\boldsymbol{\theta}, \boldsymbol{\lambda}, t)| \leq 2\Delta \quad \forall\ t \geq t_0. \tag{A.38}$$

Finally, according to assumptions of the theorem there exists a constant $\partial U \in \mathbb{R}_{>0}$ such that

$$\left|\frac{d}{dt}\upsilon_0(\hat{\boldsymbol{\lambda}}, \boldsymbol{\lambda}, t)\right| = \left|\frac{d}{dt}\boldsymbol{\theta}^T(\bar{\boldsymbol{\varphi}}_0(\mathrm{x}_0(t), \boldsymbol{\lambda}, t) - \bar{\boldsymbol{\varphi}}_0(\mathrm{x}_0(t), \hat{\boldsymbol{\lambda}}(t), t))\right| \leq \partial U$$

Hence conditions of Lemma 7 are satisfied for system (A.26) for all $t \geq t_0$. Therefore, according to Lemma 7 there exists $t_1 \geq t_0$ such that

$$|\boldsymbol{\theta}^T(\bar{\boldsymbol{\varphi}}_0(\mathrm{x}_0(t), \boldsymbol{\lambda}, t) - \bar{\boldsymbol{\varphi}}_0(\mathrm{x}_0(t), \hat{\boldsymbol{\lambda}}(t), t))| \leq 2(\sqrt{6\varepsilon\partial U} + \Delta)\ \forall\ t \geq t_1 \geq t_0 \tag{A.39}$$

provided that ε is sufficiently small. Taking (A.38) into account, we can conclude that

$$|\upsilon(\hat{\boldsymbol{\lambda}}, \boldsymbol{\lambda}, t)| \leq 2(\sqrt{6\varepsilon\partial U} + 2\Delta)\ \forall\ t \geq t_1 \geq t_0 \tag{A.40}$$

Given that (A.26) is exponentially stable and that (A.1), (A.40) hold, the upper bound of the difference $\|\hat{\boldsymbol{\theta}}(t) - \boldsymbol{\theta}\|$ at $t \to \infty$ can be estimated as follows:

$$\limsup_{t\to\infty} \|\hat{\boldsymbol{\theta}}(t) - \boldsymbol{\theta}\| \leq \frac{2}{\rho}(\sqrt{6\varepsilon\partial U} + 2\Delta).$$

Thus statement 1) of the theorem is proven.

To prove that statement 2) of the theorem holds we notice that, according to the definition of nonlinear persistency of excitation, constants L, β must exist such that for all $t \in \mathbb{R}$ there exists $t' \in [t - L, t]$:

$$|\boldsymbol{\theta}^T(\bar{\boldsymbol{\varphi}}_0(\mathrm{x}_0(t'), \boldsymbol{\lambda}, t') - \bar{\boldsymbol{\varphi}}_0(\mathrm{x}_0(t'), \hat{\boldsymbol{\lambda}}(t'), t'))| \geq \beta \cdot \mathrm{dist}(\hat{\boldsymbol{\lambda}}(t'), \mathcal{E})$$

Combining this with (A.39) yields:

$$\begin{aligned} &\forall\ t \geq t_1 \ \ \exists\ t' \in [t - L, t] : \\ &\qquad \mathrm{dist}(\hat{\boldsymbol{\lambda}}(t'), \mathcal{E}(\boldsymbol{\lambda})) \leq \frac{2}{\beta}(\sqrt{6\varepsilon\partial U} + \Delta). \end{aligned} \tag{A.41}$$

Hence there is a sequence of t'_i, $i = 1, 2, \ldots$ such that $3L \geq t'_i - t'_{i-1} \geq L$, and

$$\lim_{i\to\infty} \mathrm{dist}(\hat{\boldsymbol{\lambda}}(t'_i), \mathcal{E}(\boldsymbol{\lambda})) \leq \frac{2}{\beta}(\sqrt{6\varepsilon\partial U} + \Delta).$$

Given that $\hat{\boldsymbol{\lambda}}(t'_i)$ are bounded we can choose t'_j such that

$$\lim_{t_j\to\infty} \hat{\boldsymbol{\lambda}}(t_j) = \bar{\boldsymbol{\lambda}} \in \Omega_\lambda$$

On the other hand, according to Theorem 3 $\hat{\boldsymbol{\lambda}}(t)$ converges to a point in Ω_λ as $t \to \infty$: $\lim_{t\to\infty} \hat{\boldsymbol{\lambda}}(t) = \boldsymbol{\lambda}^*$. Therefore $\boldsymbol{\lambda}^* = \bar{\boldsymbol{\lambda}}$ and

$$\mathrm{dist}(\boldsymbol{\lambda}^*, \mathcal{E}(\boldsymbol{\lambda})) \leq \frac{2}{\beta}(\sqrt{6\varepsilon\partial U} + \Delta).$$

The latter inequality assures that

$$\limsup_{t\to\infty} \mathrm{dist}(\hat{\boldsymbol{\lambda}}(t), \mathcal{E}(\boldsymbol{\lambda})) \leq \frac{2}{\beta}(\sqrt{6\varepsilon\partial U} + \Delta).$$

Hence taking into account the bound for ε defined by (A.44), statement 2) of the theorem follows. □

A.3 Lemma 8

Lemma 8 *Consider an interconnection of systems that is governed by the following set of equations:*

$$\begin{aligned} \|\mathbf{x}(t)\| &\leq \beta(t-t_0)(\|\mathbf{x}(t_0)\| + \|\mathbf{y}(t_0)\|) + c_1 \|h(\tau)\|_{\infty,[t_0,t]} + d_1,\ d_1 \in \mathbb{R}_{\geq 0} \\ \|\mathbf{y}(t)\| &\leq \beta(t-t_0)(\|\mathbf{x}(t_0)\| + \|\mathbf{y}(t_0)\|) + c_2 \|h(\tau)\|_{\infty,[t_0,t]} + d_2,\ d_2 \in \mathbb{R}_{\geq 0} \end{aligned} \tag{A.42}$$

$$h(t) - h(t_0) = -\gamma \int_{t_0}^{t} \|\mathbf{x}(\tau)\|_\varepsilon \, d\tau,\ \gamma \in \mathbb{R}_{\geq 0} \tag{A.43}$$

where $\beta : \mathbb{R} \to \mathbb{R}$ is a strictly monotone and asymptotically decreasing to zero function. Then trajectories $\mathbf{x}(t)$, $\mathbf{y}(t)$, $h(t)$ passing through $\mathbf{x}(t_0) = \mathbf{x}_0$, $\mathbf{y}(t_0) = \mathbf{y}_0$, $h(t_0) = h_0$ at $t = t_0$ are bounded in forward time provided that

$$\gamma \leq \frac{\kappa - 1}{\kappa} \left[\beta^{-1}\left(\frac{d}{2\kappa}\right)\right]^{-1} \frac{h_0}{\beta(0)\left[\|\mathbf{x}(t_0)\| + \|\mathbf{y}(t_0)\|\right] + (c_1 + c_2) h_0 \left(1 + \frac{\kappa\beta(0)}{1-d}\right)}$$

and

$$\varepsilon \geq \left(\beta(0)\left[1 - \frac{d}{\kappa}\right]^{-1} + 1\right)(d_1 + d_2). \tag{A.44}$$

Without loss of generality we introduce a strictly decreasing sequence

$$\{\sigma_i\},\ i = 0, 1, \ldots,$$

such that $\sigma_0 = 1$, and σ_i asymptotically converge to zero. Let

$$\{t_i\},\ i = 1, \ldots$$

be an ordered sequence of time instances such that [3]

$$h(t_i) = \sigma_i h(t_0).$$

Similar to our earlier work [28] we wish to show that the amount of time needed to reach the set specified by $\|\mathbf{x}(t)\| = 0$ from a given domain of initial conditions is infinite.

The time difference $T_i = t_i - t_{i-1}$ can be estimated as

$$\begin{aligned} T_i \|\mathbf{x}(\tau)\|_{\varepsilon\infty,[t_{i-1},t_i]} &\geq \frac{h_0(\sigma_{i-1} - \sigma_i)}{\gamma} \Rightarrow \\ T_i &\geq \begin{cases} \frac{h_0(\sigma_{i-1}-\sigma_i)}{\gamma} \frac{1}{\|\mathbf{x}(\tau)\|_{\infty,[t_{i-1},t_i]} - \varepsilon}, & \|\mathbf{x}(\tau)\|_{\infty,[t_{i-1},t_i]} - \varepsilon > 0 \\ \infty, & \|\mathbf{x}(\tau)\|_{\infty,[t_{i-1},t_i]} - \varepsilon \leq 0 \end{cases} \end{aligned} \tag{A.45}$$

Suppose that $\|\mathbf{x}(\tau)\|_{\infty,[t_{i-1},t_i]} - \varepsilon \leq 0$. In this case the amount of time needed for the system to reach the set $h = h_0\sigma_i$ is infinite. Hence $h(t) \in [h_0\sigma_i, h_0\sigma_{i-1}]$ for all $t \geq t_{i-1}$, and according to (A.42), trajectories $\mathbf{x}(t)$, $\mathbf{y}(t)$, $h(t)$ will be bounded in forward time. Consider the case when $\|\mathbf{x}(\tau)\|_{\infty,[t_{i-1},t_i]} - \varepsilon > 0$ for all i.

We introduce the following sequence $\{\tau_i\}$, $\tau_i = \tau^*$, $\tau^* \in \mathbb{R}_{>0}$, $i = 1, \ldots$. Sequence $\{\tau_i = \tau^*\}$ gives rise to the series with divergent partial sums $\sum_i \tau_i = \sum_i \tau^*$. Hence proving that

$$T_i \geq \tau^* \Rightarrow T_{i+1} \geq \tau^*\ \forall\ i$$

will constitute the proof that $\mathbf{x}(t)$, $\mathbf{y}(t)$ are bounded for all $t \geq t_0$. Let $T_j \geq \tau^*$ for all $1 \leq j \leq i-1$ and consider

$$\begin{aligned} \|\mathbf{x}(\tau)\|_{\infty,[t_{i-t},t_i]} &\leq \beta(0)(\|\mathbf{x}(t_{i-1})\| + \|\mathbf{y}(t_{i-1})\|) + c_1 h_0 \sigma_{i-1} + d_1 \\ &\leq \beta(0)\left[2\beta(T_{i-1})\|\mathbf{x}(t_{i-2})\| + 2\beta(T_{i-1})\|\mathbf{y}(t_{i-2})\| + (c_1+c_2)h_0\sigma_{i-2}\right] \\ &+ (c_1+c_2)h_0\sigma_{i-1} + \beta(0)(d_1+d_2) + (d_1+d_2) \leq \\ &4\beta(0)\left[\beta^2(\tau^*)\|\mathbf{x}(t_{i-3})\| + 4\beta^2(\tau^*)\|\mathbf{y}(t_{i-3})\|\right] + P_2 \end{aligned}$$

[3] In case this equality does not hold, nothing remains to be proven; for $\|\mathbf{x}(t)\|$ will always be away from zero for all $t \geq t_0$.

where

$$
\begin{aligned}
P_2 =& \beta(0)\left[2\beta(\tau^*)(c_1+c_2)\sigma_{i-3}+(c_1+c_2)\sigma_{i-2}\right]h_0+(c_1+c_2)\sigma_{i-1}h_0 \\
&+\beta(0)\left[2\beta(\tau^*)(d_1+d_2)+(d_1+d_2)\right]+(d_1+d_2)
\end{aligned}
$$

Repeating this iteration with respect to i leads to

$$
\begin{aligned}
\|\mathbf{x}(\tau)\|_{\infty,[t_{i-t},t_i]} \leq& \beta(0)\left[2^3\beta^3(\tau^*)\|\mathbf{x}(t_{i-4})\|+2^3\beta^3(\tau^*)\|\mathbf{y}(t_{i-4})\|\right]+P_3 \\
P_3 =& (c_1+c_2)h_0\beta(0)\left[2^2\beta^2(\tau^*)\sigma_{i-4}+2\beta(\tau^*)\sigma_{i-3}+\sigma_{i-2}\right]+\sigma_{i-1}(c_1+c_2)h_0 \\
&+(d_1+d_2)\beta(0)\left[2^2\beta^2(\tau^*)+2\beta(\tau^*)+1\right]+(d_1+d_2) \\
=& (c_1+c_2)h_0\beta(0)\left[\sum_{j=0}^{2}2^j\beta^j(\tau^*)\sigma_{i-j-2}\right]+(c_1+c_2)h_0\sigma_{i-1} \\
&+(d_1+d_2)\beta(0)\left[\sum_{j=0}^{2}2^j\beta^j(\tau^*)\right]+(d_1+d_2),
\end{aligned}
$$

and after $i-1$ steps we obtain

$$
\begin{aligned}
\|\mathbf{x}(\tau)\|_{\infty,[t_{i-t},t_i]} \leq& \beta(0)(2\beta(\tau^*))^{i-1}(\|\mathbf{x}(t_0)\|+\|\mathbf{y}(t_0)\|)+P_{i-1} \\
P_{i-1} =& (c_1+c_2)h_0\beta(0)\left[\sum_{j=0}^{i-2}2^j\beta^j(\tau^*)\sigma_{i-j-2}\right]+(c_1+c_2)h_0\sigma_{i-1} \\
&+(d_1+d_2)\beta(0)\left[\sum_{j=0}^{i-2}2^j\beta^j(\tau^*)\right]+(d_1+d_2).
\end{aligned}
\tag{A.46}
$$

From (A.45) it follows that

$$
T_i \geq \frac{\sigma_{i-1}-\sigma_i}{\sigma_{i-1}}\frac{h_0}{\gamma}\frac{1}{\sigma_{i-1}^{-1}\left(\|\mathbf{x}(\tau)\|_{\infty,[t_{i-1},t_i]}-\varepsilon\right)}.
$$

Hence, if we can show that there exist $\mathbf{x}_0$, $\mathbf{y}_0$ such that such that for some $\Delta_0 \in \mathbb{R}_{\geq 0}$ and ε:

$$
\frac{\sigma_{i-1}-\sigma_i}{\sigma_{i-1}}\frac{h_0}{\tau^*} \geq \Delta_0 \tag{A.47}
$$

the following holds

$$
\gamma\sigma_{i-1}^{-1}\left(\|\mathbf{x}(\tau)\|_{\infty,[t_{i-1},t_i]}-\varepsilon\right) \leq \gamma B(\mathbf{x}_0,\mathbf{y}_0) \leq \Delta_0 \ \forall\ i \tag{A.48}
$$

where $B(\cdot)$ is a function of $\mathbf{x}_0, \mathbf{y}_0$, then boundedness of trajectories will follow.

Consider the term $\sigma_{i-1}^{-1}\|\mathbf{x}(\tau)\|_{\infty,[t_{i-1},t_i]}$, and let

$$
\sigma_i = \frac{1}{\kappa^i},\ \kappa > 1
$$

According to (A.46) we have:

$$\sigma_{i-1}^{-1}\left(\|\mathbf{x}(\tau)\|_{\infty,[t_{i-1},t_i]}-\varepsilon\right)\leq\beta(0)\left[\sigma_{i-1}^{-1}(2\beta(\tau^*))^{i-1}\right](\|\mathbf{x}(t_0)\|+\|\mathbf{y}(t_0)\|)$$
$$+\sigma_{i-1}^{-1}P_{i-1}=\beta(0)((\kappa\cdot 2\cdot\beta(\tau^*))^{i-1}(\|\mathbf{x}(t_0)\|+\|\mathbf{y}(t_0)\|)+\kappa^{i-1}P_{i-1}-\kappa^{i-1}\varepsilon$$
$$=\beta(0)((\kappa\cdot 2\cdot\beta(\tau^*))^{i-1}(\|\mathbf{x}(t_0)\|+\|\mathbf{y}(t_0)\|)+(c_1+c_2)h_0\beta(0)\kappa\left[\sum_{j=0}^{i-2}2^j\beta^j(\tau^*)\kappa^j\right]$$
$$+(c_1+c_2)h_0+\kappa^{i-1}\left(\beta(0)(d_1+d_2)\sum_{j=0}^{i-2}2^j\beta^j(\tau^*)+(d_1+d_2)-\varepsilon\right)$$

Then, choosing the value of τ^* as

$$2\kappa\beta(\tau^*)\leq d,\ \ d\in(0,1) \tag{A.49}$$

results in the following estimate:

$$\sigma_{i-1}^{-1}\left(\|\mathbf{x}(\tau)\|_{\infty,[t_{i-1},t_i]}-\varepsilon\right)\leq B(\mathbf{x}_0,\mathbf{y}_0)=\beta(0)\left[\|\mathbf{x}(t_0)\|+\|\mathbf{y}(t_0)\|\right]$$
$$+(c_1+c_2)h_0\left(1+\frac{\beta(0)\kappa}{1-d}\right)+\kappa^{i-1}\left((d_1+d_2)\left[\frac{\beta(0)}{1-\frac{d}{k}}+1\right]-\varepsilon\right).$$

Condition (A.44) implies that $\kappa^{i-1}\left((d_1+d_2)\left[\frac{\beta(0)}{1-\frac{d}{k}}+1\right]-\varepsilon\right)\leq 0$. Hence

$$\sigma_{i-1}^{-1}\left(\|\mathbf{x}(\tau)\|_{\infty,[t_{i-1},t_i]}-\varepsilon\right)\leq B(\mathbf{x}_0,\mathbf{y}_0)=\beta(0)\left[\|\mathbf{x}(t_0)\|+\|\mathbf{y}(t_0)\|\right]$$
$$+(c_1+c_2)h_0\left(1+\frac{\beta(0)\kappa}{1-d}\right)$$

Solving (A.49), (A.47) with respect to Δ_0 results in

$$\Delta_0=\frac{\kappa-1}{\kappa}\left[\beta^{-1}\left(\frac{d}{2\kappa}\right)\right]^{-1}h_0$$

This in turn implies that for all $\mathbf{x}_0$, $\mathbf{y}_0$, h_0 such that:

$$\gamma\leq\frac{\kappa-1}{\kappa}\left[\beta^{-1}\left(\frac{d}{2\kappa}\right)\right]^{-1}\frac{h_0}{\beta(0)\left[\|\mathbf{x}(t_0)\|+\|\mathbf{y}(t_0)\|\right]+(c_1+c_2)h_0\left(1+\frac{\kappa\beta(0)}{1-d}\right)}$$

the following implication must hold.

$$T_i\geq\tau^*\Rightarrow T_{i+1}\geq\tau^*.$$

Therefore, trajectories $\mathbf{x}(t)$, $\mathbf{y}(t)$, $h(t)$ passing through $\mathbf{x}(t_0)=\mathbf{x}_0$, $\mathbf{y}(t_0)=\mathbf{y}_0$, $h(t_0)=h_0$ at $t=t_0$ are bounded in forward time □

References

[1] V.I. Arnold. *Mathematical Methods in Classical Mechanics*. Springer-Verlag, 1978.

[2] G. Bastin and D. Dochain. *On-line Estimation and Adaptive Control of Bioreactors*. Elsevier, 1990.

[3] G. Bastin and M. Gevers. Stable adaptive observers for nonlinear time-varying systems. *IEEE Trans. on Automatic Control*, 33(7):650–658, 1988.

[4] A. Bobtsov, A. Pyrkin, N. Nikolaev, and O. Slita. Adaptive observer design for a chaotic Duffing system. *International Journal of Robust and Nonlinear control*, 2008. DOI: 10.1002/rnc.1354.

[5] V. I. Bykov, E. P. Volokitin, and S. A. Treskov. Parametric analysis of the mathematical model of a nonisothermal well-stirred reactor. *Combustion, Explosion and Shock Waves*, 33(3):294–300, 1997.

[6] C. Cao, A.M. Annaswamy, and A. Kojic. Parameter convergence in nonlinearly parametrized systems. *IEEE Trans. on Automatic Control*, 48(3):397–411, 2003.

[7] A. Gorban and I. Karlin. *Invarian Manifolds for Physical and Chemical Kinetics*. Lecture Notes in Physics, Springer, 2005.

[8] A.N. Gorban. *Slow relaxations and bifurcations of omega-limit sets of dynamical systems*. PhD thesis, Kuibyshev, Russia, 1980.

[9] A.N. Gorban. Singularities of transition processes in dynamical systems: Qualitative theory of critical delays electron. *Electr. J. Diff. Eqns. Monograph 5*, 2004. http://ejde.math.txstate.edu/Monographs/05/.

[10] J.L. Hindmarsh and R.M. Rose. A model of the nerve impulse using two frst-order differential equations. *Nature, Lond.*, 296:162–164, 1982.

[11] A. Ilchman. Universal adaptive stabilization of nonlinear systems. *Dynamics and Control*, (7):199–213, 1997.

[12] E. Izhikevich. *Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting*. MIT Press, 2007.

[13] G. Kreisselmeier. Adaptive obsevers with exponential rate of convergence. *IEEE Trans. Automatic Control*, AC-22:2–8, 1977.

[14] A. Loria and E. Panteley. Uniform exponential stability of linear time-varying systems: revisited. *Systems and Control Letters*, 47(1):13–24, 2003.

[15] A. Loria, E. Panteley, and H. Nijmeijer. Control of the chaotic duffing equation with uncertainty in all parameters. *IEEE Trans. on Circuits and Systems: I – Fundamental Theory and Applications*, 45(12):1252–1255, 1998.

[16] R. Marino. Adaptive observers for single output nonlinear systems. *IEEE Trans. Automatic Control*, 35(9):1054–1058, 1990.

[17] R. Marino and P. Tomei. Global adaptive output-feedback control of nonlinear systems, part I: Linear parameterization. *IEEE Trans. Automatic Control*, 38(1):17–32, 1993.

[18] R. Marino and P. Tomei. Global adaptive output-feedback control of nonlinear systems, part II: Nonlinear parameterization. *IEEE Trans. Automatic Control*, 38(1):33–48, 1993.

[19] R. Marino and P. Tomei. Adaptive observers with arbitrary exponential rate of convergence for nonlinear systems. *IEEE Trans. Automatic Control*, 40(7):1300–1304, 1995.

[20] B. Martensson. The order of any stabilizing regulator is sufficient a priori information for adaptive stabilization. *Systems and Control Letters*, 6(2):87–91, 1985.

[21] B. Martensson and J.W. Polderman. Correction and simplification to "the order of any stabilizing regulator is sufficient a priori information for adaptive stabilization". *Systems and Control Letters*, 20(6):465–470, 1993.

[22] J. Milnor. On the concept of attractor. *Commun. Math. Phys.*, 99:177–195, 1985.

[23] A. P. Morgan and K. S. Narendra. On the stability of nonautonomous differential equations $\dot{\mathbf{x}} = [\mathbf{A} + \mathbf{B}(t)]\mathbf{x}$ with skew symmetric matrix $\mathbf{B}(t)$. *SIAM J. Control and Optimization*, 37(9):1343–1354, 1992.

[24] H. Nijmeijer and H. Berghuis. On lyapunov control of the duffing equation. *IEEE Trans. on Circuits and Systems: I – Fundamental Theory and Applications*, 42(8):473–477, 1995.

[25] Jean-Baptiste Pomet. Remarks on sufficient information for adaptive nonlinear regulation. In *Proceedings of the 31-st IEEE Control and Decision Conference*, pages 1737–1739, 1992.

[26] S. Sastry and M. Bodson. *Adaptive Control: Stability, Convergense, and Robustness*. Prentice Hall, 1989.

[27] Y. Shang and B. W. Wah. Global optimization for neural network training. *Computer*, 29(3):45–54, 1996.

[28] I.Yu. Tyukin, E. Steur, H. Nijmeijer, and C. van Leeuwen. Non-uniform small-gain theorems for systems with unstable invariant sets. *SIAM Journal on Control and Optimization*, 47(2):849–882, 2008.